

Balcania et Slavia

Studi linguistici | Studies in Linguistics

e-ISSN 2785-3187

Vol. 1 – Num. 2
December 2021



Edizioni
Ca'Foscari

e-ISSN 2785-3187

Balcania et Slavia

Studi linguistici | Studies in Linguistics

Editors-in-Chief

Iliyana Krapova

Svetlana Nistratova

Luisa Ruvoletto

Giuseppina Turano

Edizioni Ca' Foscari - Digital Publishing

Fondazione Università Ca' Foscari

Dorsoduro 3246, 30123 Venezia

URL <https://edizionicafoscari.unive.it/en/edizioni4/riviste/balcania-et-slavia/>

Balcania et Slavia

Studi linguistici | Studies in Linguistics

Semestral journal

Editors-in-Chief Iliyana Krapova (Università Ca' Foscari Venezia, Italia) Svetlana Nistratova (Università Ca' Foscari Venezia, Italia) Luisa Ruvoletto (Università Ca' Foscari Venezia, Italia) Giuseppina Turano (Università Ca' Foscari Venezia, Italia)

Managing Editor Giuseppe Sofo (Università Ca' Foscari Venezia, Italia)

Advisory Board Francesco Altamari (Università della Calabria, Italia) Petya Assenova (University of Sofia, Bulgaria) Rosanna Benacchio (Università degli Studi di Padova, Italia) Stephen M. Dickey (University of Kansas, USA) Steven Franks (Indiana University Bloomington, USA) Victor A. Friedman (The University of Chicago, USA) Lucyna Gebert (Sapienza Università di Roma, Italia) Maxim Anisimovič Krongauz (Russian State University for the Humanities, Russia) Brian Joseph (Ohio State University, USA) Igor Grigorjevič Miloslavsky (Moscow State University, Russia) Vladimir Alexandrovič Plungian (Institute of Linguistics and V.V. Vinogradov Russian Language Institute; Russian Academy of Sciences; Lomonosov Moscow State University, Russia) Catherine Rudin (Wayne State College, USA) Alexander Jurijevič Rusakov (St. Petersburg State University, Russia) Andrej Nikolajevič Sobolev (Russian Academy of Sciences, St. Petersburg, Russia) Hannu Tammola (Tampere University, Finland) Björn Wiemer (Johannes Gutenberg-Universität Mainz, Deutschland)

Editorial Board Assia Assenova (Università Ca' Foscari Venezia, Italia) Pavel Duryagin (Università Ca' Foscari Venezia, Italia)

Head office Università Ca' Foscari Venezia | Dipartimento di Studi Linguistici e Culturali Comparati | Ca' Bernardo | Dorsoduro 3199, 30123 Venezia, Italia | balcaniaetslavia@unive.it

Publisher Edizioni Ca' Foscari - Digital Publishing | Fondazione Università Ca' Foscari | Dorsoduro 3246, 30123 Venezia, Italia | ecf@unive.it

© 2021 Università Ca' Foscari Venezia

© 2021 Edizioni Ca' Foscari - Digital Publishing for the present edition



Quest'opera è distribuita con Licenza Creative Commons Attribuzione 4.0 Internazionale
This work is licensed under a Creative Commons Attribution 4.0 International License



Certificazione scientifica delle Opere pubblicate da Edizioni Ca' Foscari - Digital Publishing: tutti i saggi pubblicati hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima sotto la responsabilità del Comitato scientifico della rivista. La valutazione è stata condotta in aderenza ai criteri scientifici ed editoriali di Edizioni Ca' Foscari.

Scientific certification of the works published by Edizioni Ca' Foscari - Digital Publishing: all essays published in this issue have received a favourable opinion by subject-matter experts, through an anonymous peer review process under the responsibility of the Advisory Board of the journal. The evaluations were conducted in adherence to the scientific and editorial criteria established by Edizioni Ca' Foscari.

Table of Contents

Introduction

Iliyana Krapova, Svetlana Nistratova,
Luisa Ruvoletto, Giuseppina Turano 153

From Number Cycle to Split Plurality and Back in Albanian

Gjorgje Bozhoviq 157

Theme Vowels *e* and *i* in Russian: Implementing the Cartographic Approach

Vadim Dyachkov 181

Latino o cirillico? Le proposte di un alfabeto misto nelle visioni di unità jugoslava del periodo interbellico

Giustina Selvelli 205

Кэтимологии праславянскаго обозначения ‘раны’

Artem Trofimov 225

Towards the Definition of the Bulgarian Word Order System

Anton Zimmerling 239



Introduction

Iliyana Krapova

Svetlana Nistratova

Luisa Ruvoletto

Giuseppina Turano

Università Ca' Foscari Venezia, Italia

Balcania et Slavia is happy to announce the publication of the second issue of our inaugural volume. This issue features 5 contributions in different fields of Slavic and Balkan linguistics and within different theoretical frameworks: historical linguistics, morphology, word order and clitic systems, history of the writing systems of the Slavs.

Here is a brief synopsis of the papers. All contributions are original.

In his paper, “From Number Cycle to Split Plurality and Back in Albanian”, Gjorgje Bozhoviq from the Faculty of Philology, University of Belgrade, investigates the plural morphology in Albanian. After providing a descriptive overview of the different formation strategies in modern day Albanian (affixation, stem syncretism, stem alternation, extended exponence, morphological alternances), Bozhoviq presents an analysis of Albanian plurality morphology within the Distributed Morphology Framework. The author argues that plurality in Modern Albanian can be analysed as a split number system resulting diachronically from a competition between various plural markers. Bozhoviq shows that, historically, a phonological process that reanalysed old plurals as singulars, has caused syncretism between the singular and the plural, giving rise to the creation of innovative morphological plural formations.

In his paper, “Theme Vowels *e* and *i* in Russian: Implementing the Cartographic Approach”, Vadim Dyachkov, from the Institute of Linguistics in Moscow, proposes a uniform explanation of the fact that in Russian the same thematic vowel can give rise to different structural types of verbs: the *-e* thematic vowel derives emission verbs or inchoative verbs; the *-i* thematic vowel derives causative predicates,

stative-like verbs denoting taste, stative causatives, unergative behavior-related verbs, and some others. The author proposes that, semantically, the best way to capture the differences between the two thematic vowels in Russian is to implement the Ramchand's model of predicate decomposition. He argues that it is the process sub-event that is encoded by the *-e* thematic vowel and the initial sub-event encoded by the *-i* thematic vowel. He also discusses some alternative approaches and shows the advantages of his proposal.

The paper "Latino o cirillico? Le proposte di un alfabeto misto nelle visioni di unità jugoslava del periodo interbellico" by Giustina Selvelli from the Department of Linguistic and Comparative Cultural Studies, University Ca' Foscari of Venice, discusses the context of script choice (Latin and Cyrillic) in the Kingdom of Yugoslavia in the late 1920s and early 1930s, focusing on bialphabetism and biliteracy as official policies of the country. The author places the topic in the framework of Latinization in the interwar period and examines three texts by Yugoslav authors that propose a 'hybrid' writing system containing the characters of both alphabets as a solution to digraphia. The contribution then explores some reactions to such proposals, including the one of Aleksander Belić. The article is based on the analysis of previously unknown sources found in the Matica Srpska Library in Novi Sad.

The article "On the Etymology of the Proto-Slavic Word for 'wound'" by Artem Trofimov from the Russian Presidential Academy of National Economy and Public Administration presents a critical analysis of the traditional etymology of the Proto-Slavic word with the meaning 'wound' (Old Church Slavonic рана, Greek πληγή, μάστιξ; Bulgarian pána; Serbocroatian рана, Russian рана). Linguists usually compare it to Old Indian *vraṇá-* 'wound, ulcer' and Albanian *varrë* 'wound'. Artem Trofimov offers an alternative etymology for this word proposing to revive an old idea of Per Persson, who was the first to relate the etymology of 'wound' to the Proto-Indo-European root with cognates in Proto-Slavic. Accordingly, the author gives evidence for the view that the reconstructed form of 'wound' is the Proto-Slavic **ōrna*, related to the verb **oriti* 'destroy' and the noun **ōrtъ* 'war, battle'.

In his contribution, "Towards the Definition of the Bulgarian Word Order System", Anton Zimmerling from the Pushkin State Russian Language Institute and the Institute of Linguistics Russian Academy of Science discusses the special place that Bulgarian occupies within the world's clitic systems in that it instantiates what the author labels as a W+ system. The peculiarity of the Bulgarian clitic patterns consists in the fact that this Slavic language has developed what is standardly known as an innovation with respect to the common Slavic second position cliticization pattern, namely an adjacency of the clitic to the verb such that the clitic may show up either to the left or

to the right of the verb. The author challenges this received wisdom and re-examines both synchronic and diachronic evidence in support of his claim that the Bulgarian clitic system can best be characterized as of type W+, common to some Austronesian languages, whereby the adjacency between the clitic(s) and the verb does not exclude the co-presence of (a possibly inherited) second position cliticisation. The focus of the paper are 14th-16th century texts, the Wallachian letters, which reveal an interesting tendency towards combining the two types of patterns. The author argues that, even though this can hardly be taken as the model for Modern Bulgarian, a possible historical explanation for mixed clitic behaviour might be searched in the language contacts with non-Slavic (Romance) clitic systems.

Balcania et Slavia is an online peer-reviewed journal, published annually in two issues and available in open access. The journal is published on a half-year basis by Edizioni Ca' Foscari.

We thank them for their kind collaboration and Ms Mariateresa Sala for her great editorial help. Many thanks also go to our reviewers for a constructive criticism and/or valuable suggestions for each contribution.

From Number Cycle to Split Plurality and Back in Albanian

Gjorgje Bozhoviq

University of Belgrade, Serbia

Abstract Albanian plural morphology is analysed in the paper as split between the Num and *n* heads. This is shown to result diachronically from a plural-to-singular reanalysis cycle, where splitting the number morphology between Num and *n* has worked as a competition resolution strategy in Albanian. The system is still in transition, however, and represents a long-lasting state of a dynamic equilibrium.

Keywords Albanian. Plural morphology. Split number. Reanalysis cycle. Competition resolution. Equilibrium.

Summary 1 Introduction. – 2 Albanian Plural Morphology. – 2.1 Stem Syncretism. – 2.2 Affixation. – 2.3 Stem Alternation. – 2.4 Extended Exponence. – 2.5 Other Idiosyncrasies in the Plural Formation. – 2.6 Nouns with Alternative Plurals. – 2.7 Interim Summary. – 3 Analysis. – 3.1 Split Plurality. – 3.2 Number Cycle. – 4 Discussion. – 4.1 Phonology-Driven Morphology. – 4.2 The *longue durée*. – 4.3 Equilibrium. – 5 Conclusion.



Peer review

Submitted	2021-09-13
Accepted	2021-12-03
Published	2021-12-21

Open access

© 2021 | Creative Commons Attribution 4.0 International Public License



Citation Bozhoviq, G. (2021). "From Number Cycle to Split Plurality and Back in Albanian". *Balcania et Slavia*, 1(2), [1-24] 157-180.

1 Introduction

Modern Albanian has highly elaborate and irregular plural morphology in nouns, where numerous plural markers, including affixes, stem alternation and various combinations thereof, seem to compete for the realisation of plural inflection. This produces a largely chaotic system, with many competitors and many choices, leaving the impression of a state of equilibrium, in which the morphological competition, remarkably, remains seemingly unresolved.

Several important and theoretically or typologically informing questions may be raised about such a language-particular system. First, how did it arise and how, if at all, is the competition between various plural markers eventually resolved? Secondly, how does a language-particular system like this relate to crosslinguistic typologies of plurality? And finally, what a system like this may tell us about the nature of certain morphological phenomena in language more generally? These basic questions will essentially guide this paper.

In § 2, I will provide a descriptive overview of the plural formation strategies in Albanian. These include stem syncretism (§ 2.1), affixation (§ 2.2) and stem alternation (§ 2.3), but also cases of extended exponence (e.g. simultaneous double marking) of the plural (§ 2.4), various other idiosyncratic plurals (§ 2.5) and alternative coexistent plural formations (§ 2.6).

A formal analysis of the observed data is given in § 3, where I argue, following Kramer (2016), that Albanian has a split plurality system (§ 3.1). The analysis is supplemented with a diachronic account (§ 3.2), showing how a plural-to-singular reanalysis cycle has worked to produce the present-day split plural system.

Then I will proceed to the discussion (§ 4). Although deeply grounded in a single language, in this section I will show how the analysis of Albanian plural morphology pursued here may inform our wider theoretical understanding of certain morphological phenomena, such as the dynamics of morphological change, competition resolution and states of equilibria.

2 Albanian Plural Morphology

A noun in Albanian, like *zemra* ‘heart’ in (1), can be morphologically divided into at least two parts, the grammatical stem and the inflection. The morphological shape of the stem is contingent on the noun’s gender (masculine/feminine)¹ and number values (singular/plural),

¹ Old Albanian has also inherited a third, neuter gender, which is in heavy decline in the modern language. Nouns that were originally neuter (such as *krye-t* ‘head’, *mish-të*

as in (1a) vs. (1b). *Zemra* is feminine and belongs to a small class of feminine nouns with their stem ending in an underlyingly syllabic but prosodically deficient (extrametrical for the purposes of stress assignment) *-(ë)r*. Gender assignment in Albanian nouns is both lexically and phonologically conditioned, however, not all nouns ending in *-(ë)r* will necessarily be feminine. There is also a subclass of masculines in *-(ë)r*, such as *libr-i* ‘book’ (pl. *libra-t*).

- (1)
- | | |
|--------------------------|------------------------------|
| a. <i>zemr-a</i> | b. <i>zemr:a-t</i> |
| heart. SG-NOM.DEF | heart. PL-NOM/ACC.DEF |

The inflection markers, which specify the noun’s case and definiteness, are then attached to the stem. The choice of the stem (singular or plural) and its inherent gender value in turn determines the noun’s inflection class (i.e. masculine singular, feminine singular or plural declension), as different sets of inflectional suffixes will combine only with stems of particular shape and value [table 1]. For example, a stem like *zemr-* in (1a), which is feminine singular, will combine with the feminine singular set of inflectional suffixes in table 1 (medium row), and a stem like *zemra-* in (1b), which is (feminine) plural, will combine with the plural suffixes in table 1 (the third row), when inflected for case and definiteness.

Table 1 Albanian nominal declension classes

	Indefinite paradigm				Definite paradigm			
	Nom.	Acc.	Dat.	Abl.	Nom.	Acc.	Dat.	Abl.
Masculine singular inflection				<i>-i/-u</i>	<i>-i/-u</i>	<i>-in/-un</i>		<i>-it/-ut</i>
Feminine singular inflection		–∅ (no overt marking = bare stem)		<i>-(j)e</i>	<i>-(j)a</i>	<i>-ën/-në</i>		<i>-ës/-së</i>
Plural inflection (both genders)			<i>-ve</i>	<i>-(i)sh</i>	<i>-(i)t/-të</i>		<i>-ve</i>	

‘meat’ or *ujë-t* ‘water’) are usually assigned masculine now (see e.g. Demiraj 2002, 110–11 *et passim*; Topalli 2011, 219–24). Neuter is, to a various extent, preserved in some of the (mostly Tosk Albanian) dialects (cf., e.g., Çerpja 2017 for the most recent survey, with references therein).

There are various ways to form the plural stem of a noun. In (1b), a plural suffix *-a* is apparently attached to the root *zembr-* to produce the plural form *zemra*. However, this is not the case with all plural stems, and not all plurals are even compositional like this, while some of the plurals are *overcompositional*, i.e. with more than one way of marking the plural present simultaneously. Plural formation in Albanian is a relatively complex enterprise, conditioned by the lexical, phonological and semantic features of various (sub)classes of nouns or even individual nominal roots. In addition, many nouns may form plurals in more than one possible way.²

2.1 Stem Syncretism

Several groups of nouns display number syncretism, i.e. no overt marking of the plural. This is a regular pattern in virtually all feminine nouns with stems ending in a vowel *e, i, a, o* (2), as well as in masculines formed with the agentive suffix *-s* (*-ës, -ues*) from verbs and the gentilic suffix *-as* from place names (3).

(2)

a. <i>lule: lule</i> flower.SG: flower.PL	b. <i>shami: shami</i> kerchief.SG: kerchief.PL
c. <i>kala: kala</i> fortress.SG: fortress.PL	d. <i>pallto: pallto</i> overcoat.SG: overcoat.PL

(3)

a. <i>çelës: çelës</i> key.SG: key.PL (from <i>çel</i> ‘to open’)	b. <i>mësues: mësues</i> teacher.SG: teacher.PL (from <i>mësoj</i> ‘to teach’)
c. <i>gjilanas: gjilanas</i> a person: persons from Gjiilan	

In addition to these, there are also several other nouns that manifest stem syncretism in the singular and plural lexically, i.e. as a feature of the individual nominal roots, setting them apart from the majori-

² For general synchronic surveys, cf. the standard reference works in Agalliu 2002, 94-105; Buchholz, Fiedler 1987, 249-68; or Newmark et al. 1982, 141-7. For a diachronic survey of individual plural formants in the language, see also Bokshi 2010; Demiraj 2002, 121-32; Orel 2000, 228-31 and Topalli 2011, 226-89 *et passim*, inter alia, and for the variation in plural formation across the Albanian dialects, see Fiedler 2007.

ty of nouns in their respective classes that regularly form plurals in some other way. Such are the ‘irregular’ feminines in *-ë*, which otherwise regularly form plurals by affixation, namely *gjuhë* ‘tongue(s), language(s)’, *ditë* ‘day(s)’, *rrugë* ‘road(s), street(s)’, and masculines sy ‘eye(s)’, *muaj* ‘month(s)’, *qen* ‘dog(s)’, etc.

2.2 Affixation

Affixation is by far the most elaborate plural formation strategy in Albanian. A series of different plural affixes (*-ë*, *-e*, *-a*, *-ra*, *-ër*, *-nj* etc.) compete for the plural formation in the majority of masculine nouns and in the feminine nouns with stems ending in the vowel *-ë* (schwa) or an extrametrical liquid³ *-(ë)r*, *-(ë)rr*, *-(ë)l* or *-(ë)ll*. Which plural affix will be used is mostly contingent on the lexical (e.g., gender), phonological (shape of the stem) and semantic features of a particular nominal (sub)class, so that different affixes are productive and/or sectionally restricted to a various extent.

For example, the affix *-ë* is mostly used to form plurals from masculine nouns denoting persons (4), while *-e* is mostly for inanimates (former neuters), or for masculine nouns ending in a dental (5), unless they are animate (cf. 4b). A semantic distinction is thus occasionally made in the literary language, in e.g. *element-ë* vs. *element-e* ‘elements’, where the plural formed with the suffix *-ë* would denote human entities (“representatives of a certain social stratum or class”, Kostallari et al. 1981, s.v. “element”), as in ‘foreign elements’, i.e. spies.

(4)

a. <i>shok</i> : <i>shok-ë</i>	b. <i>student</i> : <i>student-ë</i>
comrade.SG: comrade-PL	student.SG: student-PL

(5)

a. <i>send</i> : <i>send-e</i>	b. <i>mal</i> : <i>mal-e</i>
thing.SG: thing-PL	mountain.SG: mountain-PL
c. <i>shtet</i> : <i>shtet-e</i>	d. <i>fis</i> : <i>fis-e</i>
state.SG: state-PL	tribe.SG: tribe-PL

For the purposes of nominal concord, plurals formed with *-e* as in (5) pattern with feminine nouns, even though their respective singulars are masculine.

³ On the extrametricality of stem-final liquids and lexical stress assignment in Albanian in relation to its morphological structure more generally, see also Božhović 2015 and references therein.

The suffix *-a* is lexically and phonologically conditioned in the sense that it is generally found with feminine nouns ending in *-ë* (6), or with stems ending in an extrametrical liquid, irrespective of their gender (7).

(6)

a. *fushë*: *fush-a*
field.SG: field-PL

b. *vajzë*: *vajz-a*
girl.SG: girl-PL

(7)

a. *zem(ë)r*: *zemr-a*
heart.SG: heart-PL (feminine)

b. *veg(ë)l*: *vegl-a*
tool.SG: tool-PL (feminine)

c. *kumbull*: *kumbull-a*
plum.SG: plum-PL (feminine)

d. *lib(ë)r*: *libr-a*
book.SG: book-PL (masculine)

Unlike plurals in *-e*, however, the plurals in *-a* seem to be faithful to the gender of their base: all feminines that form plurals in *-a* remain feminine, and all masculines that form plurals in *-a* remain masculine in the plural for the purposes of concord.

Nouns of both genders may also form plurals with the originally collective suffix *-ra* (8). Its distribution seems to be semantically motivated to an extent, in the sense that it tends to combine with mass nouns, or with various nominal roots to produce mass readings.

(8)

a. *fshat*: *fshat-ra*
village.SG: village-PL (masculine)

b. *shi*: *shi-ra*
rain.SG: rain-PL (masculine)

c. *mall*: *mall-ra*
goods.SG: goods-PL (masculine)

d. *gjë*: *gjë-ra*
thing.SG: thing-PL (feminine)

All other plural affixes are specific to certain classes of masculine nouns. These include *-ër*, which seems to prefer some kinship terms and titles (9), *-nj* (*-inj*, *-enj*), which combines with originally nasal stems and palatalises them (10), as well as the combined suffixes *-llarë* and *-lerë*, whose first part is borrowed from the Turkish plural suffix (along with its vowel harmony) and combined with the native suffix *-ë* (11). In addition to the plural of some kinship terms and some of the historical Ottoman Turkish borrowings denoting persons high in the social hierarchy, this suffix may also be used to form family names (Demiraj 2002, 129).

(9)

a. *prind: prind-ër*
parent.SG: parent-PL

b. *mbret: mbret-ër*
king.SG: king-PL

(10)

a. *dru: dru-nj*
tree.SG: tree-PL

b. *gjarpër: gjarpër-inj*
snake.SG: snake-PL

(11)

a. *baba: baba-llarë*
father.SG: father-PL

b. *xhaxha: xhaxha-llarë*
father's brother.SG: *id.*-PL

c. *bej: bej-lerë*
Ottoman bey.SG: bey-PL

d. *Niko-llar-ë*
a toponym in southern Albania,
'descendants of Niko (Nikolla)'
(a family name turned into toponym)

There are other marginal suffixes, as well, which are attested only in several words or even in a single word. The plural of *njeri* 'human', for instance, is *njerëz*, with the plural suffix *-ëz/-zë* (Topalli 2011, 247), which is not much productive elsewhere.

2.3 Stem Alternation

A third possible way to form plural stems in Albanian, in addition to stem syncretism and affixation, is allomorphic stem alternation.⁴ Stem alternation may affect masculine nouns ending in a velar *-k*, *-g*, which palatalise to *q*, *gj* (12a-b), or a liquid *-r*, *-ll*, which palatalise to *-j* (12c-d).

(12)

a. *ujk /ujk/: ujq /ujc/*
wolf.SG: wolf.PL

b. *zog /zog/: zogj /zoɟ/*
bird.SG: bird.PL

c. *bir: bij*
son.SG: son.PL

d. *ishull /ifuɬ/: ishuj*
island.SG: island.PL

⁴ For a detailed survey of morphophonological alternations in the plural, see Memushaj (2012, 115-35).

Palatalisation, too, is largely lexically conditioned, however. Not all masculine nouns that meet the phonological conditions for palatalisation will necessarily form plurals in that way, e.g. we have *grek* ‘Greek’, with the plural form *grek-ë*, in addition to *turk* ‘Turk’, which forms the palatalised plural *turq*, patterning with *ujq* in (12a). Similarly, *mik* ‘friend’ has the plural *miq*, but *shok* ‘friend, comrade’ has a more ‘regular’ plural *shok-ë* (4a).

A related phenomenon to palatalisation, which affects consonants, is metaphony, which may affect masculine nouns with stem vowels *a* or *e*, alternating with *e* and *i*, respectively, in the plural (13). Palatalisation and metaphony often co-occur, as in (13b-d). Exceptionally, metaphony is found in a single feminine noun, as well, viz. *natë* ‘night’, which has the plural *net*.

(13)

a. <i>dash</i> /daʃ/: <i>desh</i> /deʃ/	b. <i>ka</i> /ka/: <i>qe</i> /çe/
ram.SG: ram.PL	ox.SG: ox.PL
c. <i>plak</i> /plak/: <i>pleq</i> /plec/	d. <i>mashkull</i> /maʃkuɫ/: <i>meshkuj</i> /meʃkuj/
old man.SG: id.PL	male.SG: male.PL

Historically, palatalised and metaphonic stems probably formed the plural regularly by affixation with an unattested, but reconstructable suffix **-i* (< earlier **-oi*). This segment was lost due to apocope, but it left a [+palatal] feature on the plural stems, which mutates both the non-palatal consonants and stem vowels. This mutation was allophonic at first, but later became morphophonologized (cf. Demiraj 2002, 123, 124-5 *et passim*; Topalli 2011, 231-3).

2.4 Extended Exponence

It is not uncommon for plural stems in Albanian to be marked in more than one way simultaneously, i.e. by two consecutive affixes or by stem alternation plus a plural affix. The suffixes *-llarë/lerë* in (11), for example, are one such instance of multiple exponence, to the extent to which the Turkish plural suffix *-llar/ler* is synchronically segmentable from the more productive masculine plural suffix *-ë*. A similar case is the plural of *gju* ‘knee’, *gjunjë*, which has both the palatalised nasal stem suffix *-nj* and the more ‘regular’ suffix *-ë*. A different ordering is found in *prift-ër-inj*, the plural of *prift* ‘priest’ with the nasal stem suffix following an even more idiosyncratic plural suffix *-ër*. Two subsequent plural suffixes, *-(ë)z* and *-ër*, in addition to stem-vowel metaphony, are also found in the plural form *vëlle-z-ër* from *vëlla* ‘brother’, making it arguably a unique case of triple plural marking in Albanian.

More productive cases (even with very recent loanwords) of extended exponence (i.e. double marking of the plural) include stem alternation (palatalisation, metaphony or both), in combination with a plural suffix, most frequently *-e* (14), but also *-ër* (15) and *-ra* (16).

(14)

- | | |
|---|--|
| a. <i>rrezik: rrezik-e</i>
danger, risk.SG: <i>id.</i> -PL | b. <i>bllok: bllok-e</i>
block.SG: block-PL |
| c. <i>varg: vargj-e</i>
chain.SG: chain-PL | d. <i>breg: brigj-e</i>
coast, shore.SG: <i>id.</i> -PL |
| e. <i>gardh: gjerdh-e</i>
fence.SG: fence-PL | f. <i>kopsht: qipsht-e</i>
garden.SG: garden-PL (dialectal) |
| g. <i>yll: yj-e</i>
star.SG: star-PL | |

(15)

- | | |
|--|--|
| a. <i>atë: et-ër</i>
father.SG: father-PL | b. <i>kunat: kunet-ër</i>
brother-in-law.SG: <i>id.</i> -PL |
|--|--|

(16)

- asht: esht-ra*
bone.SG: bone-PL

2.5 Other Idiosyncrasies in the Plural Formation

Not discussed above are also certain non-productive morphophonological alternations, that are characteristic only of individual roots, such as the monophthongization in the plural form *gra* from *grua* ‘woman’, or the diphthongization in *vjet* from *vit* ‘year’. These may also co-occur with palatalisation, as in *kuaj* from *kalë* ‘horse’, or take idiosyncratic plural suffixes, such as the marginal suffix *-m* in the plural form *djem* from *djalë* ‘boy’ (< **djal-m*), which also triggers stem-vowel metaphony (cf. Demiraj 2002, 130, 132; Topalli 2011, 247-8).

2.6 Nouns with Alternative Plurals

For the majority of nouns, however, plural stem formation is in fact not fixed. Many nouns can actually form more than one possible plural stem according to the patterns described above, sometimes with slightly different semantic nuances or stylistic values, as in the pair

elemente vs. *elementë* mentioned above (in § 2.2), but more often in free variation.

This is the case, e.g., for numerous feminine nouns ending in *-ë*, such as *brinjë* ‘rib’, *copë* ‘piece’, *grykë* ‘throat’, *lëkurë* ‘skin’, *pikë* ‘drop, spot’ etc., whose plurals may either be syncretic with the singular or formed regularly, by suffixing the plural marker *-a* (i.e. *brinj-a*, *cop-a* etc.). In a similar fashion, owing to its originally collective semantics, *-ra* is able to combine with most nouns that otherwise form plurals in some other way or whose plurals are otherwise syncretic (e.g. *kohë* ‘time’: plural *kohë* or *kohë-ra*; *çudi* ‘surprise, wonder’: *çu-di-ra* ‘wonders’).⁵

Similarly, many masculine nouns that meet the morphophonological criteria for stem alternation may either: (a) undergo palatalisation and/or metaphony, (b) attach an affix like *-e* in addition to stem allomorphy, thus manifesting extended exponence of the plural, or (c) form plurals with another, more ‘regular’ plural affix that does not trigger any stem alternation. This is illustrated in (17).

(17)

- | | |
|--|---|
| a. <i>oxhak</i> : <i>oxhaq-e</i> OR <i>oxhak-ë</i>
chimney.SG: chimney-PL | b. <i>hendek</i> : <i>hendetq-e</i> OR <i>hendek-ë</i>
ditch.SG: ditch-PL |
| c. <i>flamur</i> : <i>flamuj</i> OR <i>flamur-ë</i>
flag.SG: flag-PL | d. <i>portokall</i> : <i>portokaj</i> OR <i>portokall-e</i>
orange.SG: orange-PL |
| e. <i>vend</i> : <i>vend-e</i> OR <i>vis-e</i>
place.SG: place-PL | f. <i>shpend</i> : <i>shpend-e</i> OR <i>shpezë</i>
fowl.SG: fowl-PL |
| g. <i>vit</i> : <i>vjet</i> OR <i>vit-e</i>
year.SG: year-PL | |

In addition, different suffixes discussed in § 2.2 are actually in competition for the realisation of plural in many nouns, including those with (various extents of) stem alternations in addition to the affixation; cf. examples in (18).

⁵ Although *-ra* originally is a collective suffix, and the collective semantics is still felt to an extent in some of the formations, it should be noted that no difference in meaning (i.e. collective interpretation) is in fact necessary for the plurals formed with *-ra*. For the most part, *-ra* functions as a ‘regular’ plural suffix synchronically, as in *fshatra* ‘villages’ or *çudira*, with just ‘typical plural’ semantics: unlike true collective plurals, plurals in *-ra* may refer to specific groups of items, they may combine with cardinal numerals, etc. (cf. Kramer 2016, 531-2 for similar arguments from Amharic).

(18)

a. *yll*: *yj* (stem alternation) OR *yj*-e OR *yj*-ëz OR *yj*-(ë)*ra* etc.
star.SG: star-PL

b. *ujk*: *ujq* (stem alternation) OR *ujq*-ër OR *ujq*-(ë)*ra* etc.
wolf.SG: wolf-PL

c. *gjalm*(ë): *gjelm* (stem alternation) OR *gjelm*-e OR *gjelm*-ër OR *gjelmitër*
OR *gjalm*-e (without stem alternation) OR *gjalm*-a etc.
yarn.SG: yarn-PL

d. *kopsht*: *kopsht*-e OR *kop*(ë)*sht*-a OR *kopsht*-inj OR *qipsht*-e etc.
garden.SG: garden-PL

In that way, for numerous nouns, the choice between undergoing stem allomorphy or not, in addition to attaching any one of the numerous plural affixes or staying without an affix, as in (17)-(18), may even create *dozens* of different possible plural forms, which are attested in various Albanian (mostly Tosk) dialects (cf. Fiedler 2007, for a comprehensive survey of plural formation strategies in Albanian).

On the other hand, for some (mostly Geg Albanian) speakers, there is also an opposite tendency for levelling the plural variation, manifested in the expansion of the affix *-a* to create forms like *element-a*, *dokument-a*, *plakat-a*, *autobus-a*, *student-a*, *oxhak-a* (cf. in contrast 17a), *çelës-a* (cf. the unaffixed plural *çelës* in (3) above), etc. Recall that *-a* is the most frequent plural suffix for feminine nouns, and that it also combines with a subclass of masculine stems. Analogically from there, it spreads to encompass the majority of other masculines as well, thus eliminating both the other plural affixes that are in competition with each other, and the stem allomorphy that is associated with them and that regularly produces cases of multiple exponence (double marking) of the plural.

2.7 Interim Summary

So far we have seen how the plural stem formation strategies in Albanian may vary from stem syncretism (i.e. no overt marking of number whatsoever), via more or less productive affixation, with more or less selectional restrictions on particular plural affixes, to various stem allomorphy patterns and purely idiosyncratic formations. Particularly curious are instances of extended exponence (including apparent borrowing of an inflectional affix from Turkish), and an almost unhindered variation and irregularity in Albanian nominal plural formation.

In the following section, I will provide a formal analysis of the observed competition in plural marking in Albanian, building mostly

on Kramer (2016), but supplying it also with a diachronic account of how (and why) this dynamic plural formation system came into place.

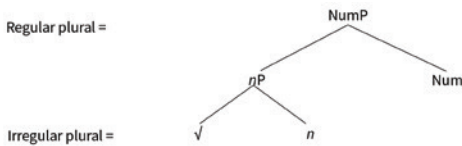
3 Analysis

Since Ritter (1991), it has been argued for the existence of a syntactic node Num(ber), the head of an additional functional projection Num(ber)P above the lexical NP (between DP and NP), to accommodate number inflection. NumP is assumed to be crosslinguistically the default syntactic location of number features within the DP. On the other hand, more recent research on idiosyncratic plurality, such as irregular and ‘lexical’ plurals, argues for a possibility that number features may also be found elsewhere in the DP, most notably on the categorial head *n*, i.e. closer to the root (cf. Acquaviva 2008; Alexiadou 2011; Harbour 2011, among others; Kramer 2016 for a survey and references therein).

Assuming a DM-style word structure, in which lexical categories are made up of a category-neutral root (*v*) and a category-determining head, as the little *n* for nominals (cf. e.g. Marantz 1997; 2001), Kramer argues that number features are in fact split between Num for morphologically regular plurals and the nominalizing head *n* for idiosyncratic (i.e. irregular) plurals in Amharic (19), and proposes a three-way morphosyntactic typology of plurality, as in (20), in which:

[l]anguages can have wholly Num-based plurals (e.g. English), *n*-based plurals (or possibly closer-to-root plurals, where all plurals act idiosyncratically), or split plurality (an irregular/regular contrast in plurality expressed through two different heads). (Kramer 2016, 555)

(19)



(adapted from Kramer 2016, 534)

(20)

Morphosyntactic typology of plurality (Kramer 2016, 555):

- a. Num-based plurality (English, Hebrew, etc.)
- b. *n*-based plurality (Somali, Halkomelem Salish, etc.)
- c. Split plurality (Amharic)

In the remainder of this section, I will argue that the Albanian plural formation system belongs to the split or ‘hybrid’ type, involving both Num and *n* heads, like that of Amharic (20c), according to most of the diagnostics for split plurality that are identified in Kramer (2016). After that, I will show how this particular formal analysis relates to the facts of Albanian historical grammar. In particular, I will argue that the phenomenon described in classical Albanology as “singularization of plurals” (Çabej 2012) can be formally modelled as a Jespersen-like plural-to-singular reanalysis cycle,⁶ in which the reanalysis of old plurals as singulars has led to a ‘renewal’ of various innovative plural forms. In formal terms, it will be shown that the old plural morphology was ‘pushed down’ from Num to *n* and reanalysed as part of the nominal stem, so that various innovative plural markers could appear in (now empty) Num, competing among themselves for the realisation of plural morphology. A “niche competition” was created, in the sense of Aronoff (2016; 2019), so that, in effect, splitting the plural features between Num and *n*, in order to maintain a system like the one in (19), has emerged in Albanian as a possible competition resolution strategy.⁷

3.1 Split Plurality

A crucial notion for understanding split plurality systems is that of morphosyntactic *regularity* of a plural formation strategy, which is reflected in its productivity and the degree of selectional restrictions. ‘Regular’ morphology tends to be more productive, with less selectional restrictions, while ‘irregular’ morphology entails more lexically constrained and idiosyncratic structures, which are therefore less productive.

Albanian plural formation strategies, as described in § 2, may be represented on a continuum of morphosyntactic regularity along those lines, as in **table 2**.

⁶ Jespersen (1917) is credited for formulating a cyclic change in sentential negation marking, which is now termed *Jespersen’s cycle*. In a cyclic change like this, historical loss or weakening of a grammatical expression is followed by its ‘renewal’ by another element in the same function. Many such cyclic changes, involving not only negation, but other functional items as well, including pronouns, auxiliaries and modals, prepositions etc., have been described since (e.g. cf. van Gelderen 2009; 2011; 2016).

⁷ For a similar connection between a Jespersen-like cycle and competition resolution, which is advocated for in demonstratives, see also Vindenes 2018.

Table 2 The continuum of morphosyntactic regularity in Albanian nominal plurals

syncretism	more ‘regular’ (i.e. productive) plural affixes (like -a, -e, -ë)	more lexically restricted affixes (like -ër, -llarë, -ëz, etc.)	stem allomorphy	doubly marked and other lexically idiosyncratic plurals
more regular, least selectional restrictions ← → less regular, most selectional restrictions				

It should be noted that not every plural system in which different morphological strategies or vocabulary items compete for the realisation of plural is a split one, however. On the contrary, the mere existence of regular and irregular plural morphology in the context of different lexical roots, is rather an argument (as per Occam’s razor) in favour of plural features hosted under the *single* Num node. In systems like these, the grammar resolves the competition by imposing contextual restrictions on different allomorphs that compete for the realisation of Num. In such cases, the competition is resolved by the Pāṇinian (‘elsewhere’) principle, and such systems are therefore *not* split systems.

In split number systems, on the other hand, the difference between regular and irregular morphology is a structural (formal) one, not just realisational. Here, regular plural morphology is inserted at Num[+PL], while the high selectional restrictions and lower productivity of the ‘irregular’ plurals come instead from their being of an entirely different category; namely, the realisations of plural features under the lower head *n*. This is why double (or even *triple*) plural marking, with both ‘regular’ and ‘irregular’ markers, may co-occur on a single noun (recall § 2.4), a thing that should be impossible under the Pāṇinian principle, but is readily available in Albanian. Also, this is why there are overabundant alternative ways to form plurals even from the same noun in Albanian (as in § 2.6), either using the more ‘regular’ or the ‘irregular’ morphological strategies from table 2, or a combination thereof. Again, under the Pāṇinian principle, where the competition is contextually resolved, that too should be impossible. In other words, in split systems like these, as Kramer (2016) points out, the regular and irregular morphology do not actually compete for the realisation of the same syntactic head; rather, they are associated with separate functional projections (as in 19).

The Albanian situation as sketched in table 2 is, *mutatis mutandis*, highly reminiscent of the one described by Kramer for Amharic, where the morphologically ‘regular’ plurals are formed via a Num[+PL] combining with a *nP*, whereas ‘irregular’ plurals are in fact formed by combining *n*[+PL] morphology with the root. In order to test the hypothesis that Albanian is a split-number language like Amharic, we will check the Albanian data against the diagnostics that are used as evidence for a Num/*n* split by Kramer (2016, 532–40). These are summarised in (21).

(21)

Evidence for a Num/*n* split (from Kramer 2016, 540):

- a. Ordering of plural morphemes in the double plural
- b. Lack of productivity of the irregular plural vs. uniform productivity of the regular plural
- c. Selectional restrictions on the irregular plural vs. lack of restrictions for the regular plural
- d. Semantic idiosyncrasies more common for the irregular plural
- e. Gender distinctions displayed only by irregular plurals
- f. Ban on realizing a gender suffix and an irregular plural suffix at the same time
- g. Category changes accomplished only by irregular plurals

As for the most obvious argument from morpheme ordering (21a), recall that in cases of simultaneous multiple marking of the plural, the more ‘irregular’ a plural formation strategy is on the continuum in table 2, the closer it will appear to the root (or even surface *on* the root itself, in the case of stem alternations). This was the case, e.g., with *gju-nj-ë* ‘knees’, *baba-llar-ë* ‘fathers’ etc., where the more productive, less idiosyncratic masculine plural suffix *-ë* comes after the more lexically restricted affixes *-nj* or *-llar*. Given that affixes *-nj* and *-llar* are stem-formative (for the nasal stems) and subcategorizing (defining a highly specific subclass of masculine nouns), respectively, they are more reminiscent of the nominalising morphology that we expect to find at *n*, while the distributional properties of *-ë* are more like the ‘typical’ plural inflectional morphology at Num.

More ‘irregular’ plural formation strategies in table 2 are *relative-*ly less productive and more restricted than the more ‘regular’ formations, as per (21b-c). Unlike Amharic, however, there is not a sharp contrast here, so that not even the most regular plurals are uniformly productive in Albanian. This may be accounted for by the historical and typological differences between the two languages: namely, Albanian has inherited gender-based lexical restrictions for all its inflectional markers from the proto-language, because that is how the competition between different affixes was resolved typically for Indo-European languages at an earlier stage, in that way still ‘feeding’ their present-day distribution in the language. Nevertheless, there is a clear continuum in the *degree* of productivity and the amount of selectional restrictions between different plural markers in Albanian, so that e.g. an affix like *-a* may combine with both masculines and feminines, irrespective of their gender (recall, e.g., feminine *zembr-a* and masculine *libr-a*), while there are other affixes or plural formation strategies such as stem allomorphy that only affect masculine stems or even specific subclasses of masculines (recall § 2.3 and purely lexical discrepancies such as the plurals *grek-ë* from *grek* ‘Greek’ vs. *turq* from *turk* ‘Turk’, or *shok-ë* from *shok* ‘friend, comrade’ vs. *miq* from *mik* ‘friend’).

This is what makes the irregular plurals in a split plurality system more ‘derivation-like’ than ‘inflection-like’, according to Kramer, which is precisely because of their belonging to *n*-related, not Num-related morphology:

the ‘irregular’ plural is only attested with certain roots, and paradigmatic gaps are typical of category-determining head (e.g., *n*) and root combinations [...] In other words, irregular plurals have selectional restrictions on stems. This is highly reminiscent of (traditionally defined) derivational morphology, which also has strict selectional restrictions [...] Although the derivational / inflectional distinction has no theoretical status in Distributed Morphology, it often corresponds to the distinction between non-category-defining heads like Num (inflectional) and category-defining heads like *n* (derivational). (Kramer 2016, 535)

In Albanian, this is additionally corroborated by the fact that the more ‘irregular’ plural marking is typically preserved in derivation, e.g. in the adjective formed with *-or*, *prindër-or* ‘parental’, from *prind* (9); the adverb formed with *-isht*, *ujqër-isht* ‘wolf-like’, from *ujk* (12 viz. 18b); the feminine formed with *-eshë*, *bejler-eshë* ‘wife of a bey’, from *bej* (cf. 11), etc.

The higher sensitivity of the more ‘irregular’ plurals in Albanian to the gender of the stem corresponds, at the same time, to Kramer’s criteria (21e-f), although again, allowing for the apparent typological differences between Albanian and Amharic. As we already noted, the more ‘irregular’ plural markers in Albanian are inherently gendered, while the more ‘regular’ markers like *-a* may equally take both masculine and feminine stems. According to Kramer,

[t]he restriction of gendered plurals to irregular plurals is puzzling unless gender is a feature on *n*, thus creating a feature bundle that has both gender and number. (2016, 537)

Recall also that, in Albanian, some plural affixes like *-e* may switch the gender of the noun from masculine to feminine (so that the singulars like *send*, *mal*, *shtet* are masculine for the purposes of concord, while the respective plurals *send-e*, *mal-e*, *shtet-e* are feminine; cf. *Shtetet e Bashkuar-a* ‘United-FEM States’).⁸

A split analysis of plurality also predicts that idiosyncratic interpretations will be more likely with the plurals associated with *n*, than with the ‘regular’ plurals associated with Num (21d). Also, by virtue

⁸ For more on this phenomenon from a diachronic and a Balkanological perspective, see Demiraj 1966; Topalli 1995, *inter alia*.

of being categorising heads, the plural markers associated with *n* may serve as nominalising affixes, while those associated with Num may not engage in category changing (21g). Although this is not observed in our Albanian data, this is readily explainable, too, by the fact that semantics and derivational processes are to a large extent language-specific, and that not all formal possibilities may be necessarily exploited in every language. This is not to say that semantic idiosyncrasies or category changes with the more ‘irregular’ plural morphology in Albanian are impossible. Rather, the Num/*n* split in Albanian plural morphology should be understood as a more *gradual* one, in which the plural markers that may be associated with *n* have been relatively recently ‘pushed down’ from Num to *n* by reanalysis, and this is a still ongoing process. Therefore, due to the historical ‘noise’ and the differences in the exact details of their respective grammaticalization paths, it is not perhaps as straightforward as it is in Amharic. The exact dynamics of this reanalysis cycle will be discussed in the following section.

3.2 Number Cycle

A split analysis of plurality in Albanian is also corroborated diachronically. Albanian is shown to have undergone a process called “singularization of plurals” (cf., e.g., Çabej 2012; Demiraj 1985; 2002, 119–21; Orel 2000, 223–4 et passim; Topalli 2011, 289–93), by which the historical plural forms were reanalysed as singulars, and the original singular forms of nouns were lost. This has led to a phase of formal syncretism between the singular and the plural, which then sparked the creation of innovative plural formations, in order to restore the formally lost number opposition.

This process is attested in its phase of plural-to-singular syncretism in Old Albanian authors, e.g. in Buzuku we find *qytet* as both the plural and the singular form of the present-day *qytet* ‘town’, plural *qytet-e* (Demiraj 2002, 119–20). Being a borrowing from the Latin *civitat-em*, the phonetically regular Albanian reflex should have been **qytat* in the singular. The attested form *qytet* is apparently a metaphonized plural of **qytat*, which was in Buzuku’s times ‘singularised’, i.e. reanalysed as a singular. An innovative plural formation *qytet-e*, with a ‘redundant’ plural affix *-e*, has entered the language in the meantime, in order to restore the formal opposition between the singular and the plural.

A similar plural-to-singular reanalysis cycle is described by Bates and McKenzie (2021) for pronominal systems. It is shown to proceed through five phases, initiated by semantic reanalysis (feature drop), whereby the originally plural form β , in the pair $\alpha_{[sc]} : \beta_{[pl]}$, loses its plural feature and becomes a general form (‘bare β wins’), while the

singular α is lost. In some languages, then, the cycle may continue when a new plural form $\gamma_{[\text{PL}]}$ emerges to denote the plural and pushes the general form to denote the singular (‘singularise’). The number opposition in $\beta_{[\text{SG}]} : \gamma_{[\text{PL}]}$ is thus reinstalled **[fig. 1]**.

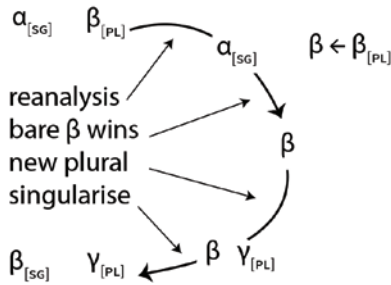
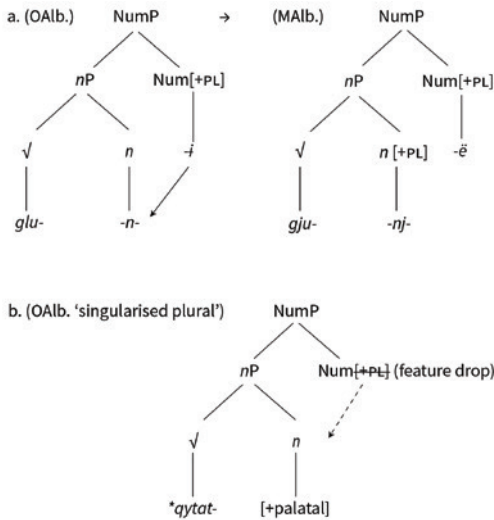


Figure 1 The plural-to-singular reanalysis cycle (adapted from Bates, McKenzie 2021, 13)

Extending this analysis to the nominal domain, we may observe all five stages of the number cycle at play in Albanian. The initial stage had a morphologically regular, unattested but reconstructable opposition in number marking, e.g. **qytat* : **qytat-i*. The original inflection **-i* was lost due to apocope, but left a trace in the initially allophonic stem-vowel umlaut (**qytät-i* > **qytät-*). This form was then phonologized and generalised as Buzuku's syncretic ('bare') form *qytet*. A new affixal plural *qytet-e* emerges and the old plural *qytet* is now singularised.

A split plurality analysis, advocated in the previous section, equally accounts for the historical processes of the ‘singularization of plurals’ (i.e. loss of the original number distinction) and the emergence of innovative plural markers. The loss of old plural markers such as **-i* and the consequent semantic reanalysis and generalisation of the old plural stems, has left the Num node morphologically empty, either by literally deleting the segmental material associated with it, or by ‘pushing down’ the old plural morphology, synchronically surviving as stem allomorphy, from Num to *n*, in analogy with the now prevalent apocopated plurals. This has allowed for other (more ‘regular’ from a synchronic point of view) affixal plural markers to redundantly reappear and to compete among themselves for the realisation of Num. This has also produced the possibility of double marking of the plural, as in *gju-nj-ě*, where the nasal stem formant is both palatalized, thus preserving trace of the old plural marker that was apocopated (viz. lowered to *n* by reanalysis), and followed by an innovative plural affix *-ě*. This is illustrated in (22a). An intermediate (i.e. Buzuku) stage with generalised plurals is shown in (22b).

(22)



The loss of the original number distinction, i.e. lowering of Num morpho(phono)logy to *n*, has created a very unstable system in Albanian, in which speakers have been forced to employ numerous available plural formation strategies in order to reinstall the opposition, including different plural affixes and stem allomorphy (and various possible combinations thereof). This is the reason behind the almost unhindered variation in Modern Albanian plural stem formation, as no particular morphological strategy has had an advantage over another. In that sense, splitting the number features between Num and *n* has worked as a competition resolution strategy in Albanian: different plural formation strategies are associated with separate syntactic heads and therefore won't compete for the realization of the same node.

However, the competition resolution in Albanian plural morphology is still an ongoing process, and this is why some plural markers (like *-ër*, which in some cases is reminiscent of 'regular' plurals, and in others seems closer to 'irregular' plurality) may occasionally fluctuate between both Num and *n*. Consider, e.g., the extended plurals *vëllezër* 'brothers' and *priftërinj* 'priests'. In the former, *-ër* is the second suffix (after the more restricted *-ëz*), and represents a regular plural formation strategy for masculine kinship terms. In the latter, however, *-ër* is the middle suffix, and it is more restricted and less productive in comparison to the more regular plural extension (for nasal stems) *-inj*. In this case, *-ër* effectively functions as a categorising morpheme, which is further corroborated by it appearing in other nouns derived from *prift*, e.g. in the feminine *priftër-eshë* 'priestess; priest's wife'. This is why a continuum like the one in ta-

ble 2, rather than sharp categorical distinctions, is better suited for the description of such fluctuating and fuzzy phenomena.

Another possible option for resolving the ongoing competition between various plural markers in Albanian is analogical levelling, and this is precisely what speakers of some dialects opt for, by expanding the use of the affix *-a* (recall, e.g., *element-a*, *oxhak-a*, *çelës-a* etc. in § 2.6). Both competition resolution strategies – splitting the plural markers between different heads and levelling a single marker across all stems – are, as we have seen, ultimately motivated by the loss of the original Num(ber) morphology.

4 Discussion

In this section, I will briefly comment on the broader theoretical significance of the previously analysed Albanian data. There are three main points that the Albanian number cycle and split plurality allow us to make, and these are: (a) the role of phonology in morphological change, (b) the durativity of mechanisms driving structural change, and (c) the dynamics of competition resolution and the nature of morphological equilibria.

4.1 Phonology-Driven Morphology

The plural-to-singular reanalysis cycle, like the one we have described for Albanian, is argued to be motivated by semantic reanalysis (feature drop) and overall pragmatic factors in the pronominal systems (Bates, McKenzie 2021). Likewise, in Albanian historical grammar (following the analysis made in Çabej 2012), the cause of the ‘singularisation of plurals’ is also often sought in semantics, viz. the inherent “intensifying” or “collectivizing” (thus e.g. Bokshi 2005) meaning of the plural forms, which allows for their generic use. Other views include speculations on the possible frequency effects and the formal factors driving the plural-to-singular reanalysis, such as phonological change and the working of analogy (Demiraj 2002, 119-20; Topalli 2011, 290-2).

However, given that the number cycle in Albanian was formally initiated by the apocope of word-final syllables, which contained the inflectional material, one has to account for the role of phonology in morphological change. Even though the question of primacy of form or meaning in language is always one of ‘the chicken and the egg’, and leaving aside for now the speculations regarding the ultimate cause and motivation of the diachronic ‘singularization of plurals’ in Albanian, what this particular plural-to-singular reanalysis process shows is that at least in the sequence of events, phonological change precedes, if not causes, morphological reanalysis. In this case, both

semantic reanalysis with the generalisation of plural stems, and the subsequent emergence of innovated plurals, were preceded and may very well be motivated by the fact that the original morphological information was simply lost due to a regular phonological process.

The decisive effect of phonological processes is visible in the synchronically surviving stem alternations in Albanian. Following the loss of old plural markers, an originally allophonic variation of the stem consonants and vowels was morpho(phono)logized as stem allomorphy. From a diachronic perspective, this conforms to some autonomous approaches to morphology, according to which “morphology exists because morphophonology exists”, i.e. “phonology gives rise to morphophonology which gives rise to morphology” (Carstairs-McCarthy 2010).

4.2 The *longue durée*

Once a morpho/phonological change is on its way, as has happened when the loss of old plural marking initiated the number cycle and plural split in Albanian, it may have long-lasting effects on language structure. E.g. splitting number features in Albanian was not a single event in which various plural markers were distributed between Num and *n* at once, but it has had a lasting, profound influence on various plural formations and related phenomena in the language.

Long-lasting processes like these usually produce fluctuating, gradual continua with fuzzy boundaries, rather than neat and sharp categorical contrasts. Crosslinguistically, this is why many language-particular systems will not fit easily into some of the theoretically predefined ‘types’, as the ones in (20). For this reason, a plural system like Albanian may be better described in diachronic terms, as *transitory* between an (older) Num-based plurality and a (more innovative) split plurality.

At the same time, however, every single change in the language creates pressure on the system to resolve and re-establish regularity. This is why all diachronic changes ultimately *are* directional toward some (system-internal) goal.

4.3 Equilibrium

Finally, while it is true that speakers will aim to (eventually) resolve an ongoing competition in the language, and that there are, in that sense, no true stable equilibria (as per Gause’s ecological principle; cf. Aronoff 2019), competition resolution itself may be a long-lasting process, not a single event, so that at any given time it may still have no clear outcome, and hypothetically may even ultimately fail.

This is currently the case with the Albanian plural system. A seemingly naïve (morpho)phonological split between the historical and emerging plurals in the language has led to an apparent state of equilibrium, in the sense of Aronoff (2016, 47), “in which no competitor has a clear advantage and such states may persist for long periods”.⁹ Yet, this state of equilibrium is not static, as different competition resolution strategies, including Pāṇinian-style blocking of certain plural affixes, analogical levelling and the general shift toward a split-number system, have been operating in Albanian in an attempt to resolve the competition induced by the number cycle.

Still, this dynamic equilibrium is best manifested in the abundance of alternative plural formations for the majority of nouns, including the plural forms that have been innovatively (re-)produced by the aforementioned competition resolution strategies themselves, thus *expanding* the stock of available plural formations, rather than reducing it as would be expected.

5 Conclusion

In this paper, I have argued that plurality in Modern Albanian is best analysed as a split number system resulting from an ongoing competition between various plural markers, that was in turn triggered by the phonological history and reanalysis of historical plurals in the language.

The analysis, based on Kramer (2016), contributes the Albanian data to the existing corpus of literature on distributed number morphology across different functional heads in the nominal domain.¹⁰ In addition, supplemented with a language-specific diachronic perspective, it emphasises the gradual and transitory nature of the N/n split in Albanian.

In turn, this may add to our better understanding of the structural mechanisms behind the inception, duration and the ultimate outcome of morphological competition and the states of equilibria.

⁹ Aronoff makes note of a similar state of equilibrium in Maltese plural morphology.

¹⁰ See, e.g., most recently Dali and Mathieu 2021 for a monographic treatment focusing mostly on Arabic; Božović 2020 for a similar preliminary analysis of the West South Slavic data.

Bibliography

- Acquaviva, P. (2008). *Lexical Plurals*. Oxford: Oxford University Press.
- Agalliu, F. (2002). "Emri" (Noun). Demiraj, Sh. (ed.), *Gramatika e gjuhës shqipe* (Grammar of the Albanian Language). Vol. 1, *Morfologjia* (Morphology). Tiranë: Akademia e Shkencave e Shqipërisë, 81-152.
- Alexiadou, A. (2011). "Plural Mass Nouns and the Morpho-syntax of Number". Byram Washburn, M. et al. (eds), *WCCFL 28: Proceedings of the 28th West Coast Conference on Formal Linguistics*. Somerville (MA): Cascadia Press, 33-41.
- Aronoff, M. (2016). "Competition and the Lexicon". Elia, A. et al. (eds), *Livelli di analisi e fenomeni di interfaccia = Atti del XLVII congresso internazionale della società di linguistica italiana*. Roma: Bulzoni Editore, 39-52.
- Aronoff, M. (2019). "Competitors and Alternants in Linguistic Morphology". Rainer, F. et al. (eds), *Competition in Inflection and Word-Formation*. Cham: Springer, 39-66. https://doi.org/10.1007/978-3-030-02550-2_2.
- Bates, J.; McKenzie, A. (2021). "A Plural-to-Singular Reanalysis Cycle". *Journal of Historical Syntax*, vol. 5, article 24. <https://doi.org/10.18148/hs/2021.v5i16-25.65>.
- Bokshi, B. (2005). *Rruga e formimit të fleksionit të sotëm nominal të shqipes* (The Formation Path of the Modern Albanian Nominal Inflection). Prishtinë: Akademia e Shkencave dhe e Arteve e Kosovës.
- Bokshi, B. (2010). *Periodizimi i ndryshimeve morfologjike të shqipes* (Periodisation of Morphological Changes in Albanian). Prishtinë: Akademia e Shkencave dhe e Arteve e Kosovës.
- Božović, Đ. (2015). "Fonološko-gramatički interfejs slogovnog akcenta" (The Phonology-Grammar Interface in Lexical Stress). *Anali Filološkog fakulteta* (Annals of the Faculty of Philology), 27(2), 369-401. <https://doi.org/10.18485/anaLiff.2015.27.2.20>.
- Božović, Đ. (2020). "Dual Nature of Dual Number in Western South Slavic". Audring, J.; Koutsoukos, N.; Manoulidou, C. (eds), *Rules, Patterns, Schemas and Analogy = Proceedings of the 12th Mediterranean Morphology Meeting*. University of Patras, 18-27. <https://doi.org/10.26220/mmm.3246>.
- Buchholz, O.; Fiedler, W. (1987). *Albanische Grammatik* (Albanian Grammar). Leipzig: Verlag Enzyklopädie.
- Çabej, E. (2012). *Shumësi i singularizuar në gjuhën shqipe* (Singularised Plural in the Albanian Language). Tiranë: Botime Çabej.
- Çerpja, A. (2017). "Çintegrimi i asnjansit dhe mbetje të tij në të folme të shqipes" (Disintegration of the Neuter and Its Remnants in Albanian Dialects). *Seminari Ndërkombëtar për Gjuhën, Letërsinë dhe Kulturën Shqiptare* (International Seminar on Albanian Language, Literature, and Culture), 35, 145-56.
- Carstairs-McCarthy, A. (2010). *The Evolution of Morphology*. Oxford: Oxford University Press.
- Dali, M.; Mathieu, E. (2021). *A Theory of Distributed Number*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.269>.
- Demiraj, S. (1966). "Rreth ndërrimit të gjinisë së emrave në shumës në shqipe vështuar edhe në përjasje me rumanishten" (On Gender Switch in Plural Nouns in Albanian Also Investigated in a Comparison with Romanian). *Studia filologjike* (Philological Studies), 20(4), 83-97.

- Demiraj, S. (1985). "Singularizimi i shumësit të emrave në gjuhën shqipe (vështrim historik)" [Singularization of Plurals in Albanian Nouns (A Historical Study)]. *Studime filologjike* (Philological Studies), 39(4), 13-22.
- Demiraj, S. (2002). *Gramatikë historike e gjuhës shqipe* (A Historical Grammar of the Albanian Language). Tiranë: Akademia e Shkencave e Shqipërisë.
- Fiedler, W. (2007). *Die Pluralbildung im Albanischen* (The Formation of Plural in Albanian). Prishtinë: Akademia e Shkencave dhe e Arteve e Kosovës.
- Harbour, D. (2011). "Valence and Atomic Number". *Linguistic Inquiry*, 42(4), 561-94. https://doi.org/10.1162/LING_a_00061.
- Jespersen, O. (1917). *Negation in English and Other Languages*. København: Høst & Son.
- Kostallari, A. et al. (1981). *Fjalor i gjuhës së sotme shqipe* (Dictionary of the Contemporary Albanian Language). Prishtinë: Rilindja.
- Kramer, R. (2016). "A Split Analysis of Plurality: Number in Amharic". *Linguistic Inquiry*, 47(3), 527-59. https://doi.org/10.1162/LING_a_00220.
- Marantz, A. (1997). "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon". Dimitriadis, A. et al. (eds), *Proceedings of the 21st Annual Penn Linguistics Colloquium*. Pennsylvania Working Papers in Linguistics, 4(2), 201-25.
- Marantz, A. (2001). *Words* [unpublished manuscript]. MIT.
- Memushaj, R. (2012). *Morfonologji e eptimit të shqipes standarde* (Morphophology of Inflection in Standard Albanian). Tiranë: Infbotues.
- Newmark, L. et al. (1982). *Standard Albanian: A Reference Grammar for Students*. Stanford (CA): Stanford University Press.
- Orel, V. (2000). *A Concise Historical Grammar of the Albanian Language: Reconstruction of Proto-Albanian*. Leiden; Boston; Köln: Brill. <https://doi.org/10.1163/9789047400363>.
- Ritter, E. (1991). "Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew". Rothstein, S.D. (ed.), *Perspectives on Phrase Structure: Heads and Licensing*. San Diego: Academic Press, 37-62. https://doi.org/10.1163/9789004373198_004.
- Topalli, K. (1995). "Rreth një paradigme të zhdukur të emrave të gjinisë femërore" (On One Lost Paradigm of Feminine Nouns). *Studime filologjike* (Philological Studies), 49, 15-24.
- Topalli, K. (2011). *Gramatikë historike e gjuhës shqipe* (A Historical Grammar of the Albanian Language). Tiranë: Botimet Albanologjike.
- van Gelderen, E. (ed.) (2009). *Cyclical Change*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.146>.
- van Gelderen, E. (2011). *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199756056.001.0001>.
- van Gelderen, E. (ed.) (2016). *Cyclical Change Continued*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.227>.
- Vindenes, U. (2018). "Cyclic Renewal of Demonstratives". *Studies in Language* 42(3), 641-68. <https://doi.org/10.1075/sl.17047.vin>.

Theme Vowels *e* and *i* in Russian: Implementing the Cartographic Approach

Vadim Dyachkov

Institute of Linguistics, RAS, Moscow

Abstract The paper deals with the Russian theme vowels that can be regarded as derivational suffixes, in particular, with *e*- and *i*-vowels. The former derives emission verbs or inchoative verbs, whereas the latter derives causatives (both dynamic and stative), unergative behaviour-related verbs, and some others. I propose a uniform explanation of the fact that the same theme vowel can give rise to different structural types of verbs. I argue that, semantically, the best way to capture the differences between the two vowels is to implement the Ramchand's model of predicate decomposition, where *e*-vowel encodes the process sub-event and *i*-vowel the initial sub-event.

Keywords Russian. Theme vowels. Unaccusativity. Unergativity. Denominal verbs. Deadjectival verbs. Stative causatives.

Summary 1 Introduction. – 2 Structural Types of Denominal and Deadjectival Verbs. – 2.1 *e*-Verbs. – 2.2 *i*-Verbs. – 3 Formal Implementation in Ramchand's Model. – 4 Structural Types of *i*-Verbs. – 4.1 Stative Causatives. – 4.2 Predicates of Taste. – 4.3 Instrumental Verbs. – 4.4 Unergative Verbs. – 4.5 Relations Between *e*- and *i*-Verbs. – 5 Discussion. – 6 Conclusion.



Peer review

Submitted	2021-07-20
Accepted	2021-12-14
Published	2022-02-15

Open access

© 2021 | Creative Commons Attribution 4.0 International Public License



Citation Dyachkov, V. (2021). "Theme Vowels *e* and *i* in Russian: Implementing the Cartographic Approach". *Balcania et Slavia*, 1(2), 181-204.

1 Introduction

The fact that some Slavic theme vowels are elements going back to derivational affixes is well-known in historical linguistics (cf. Lunt 2001; Vaillant 1966). However, the same theme vowel can give rise to several structural and/or semantic types of verbs. In this paper, I will focus on two theme vowels deriving denominal and deadjectival verbs, namely *-e* and *-i*, and explore their semantic and structural properties.

There is evidence that at least some of Slavic theme vowels are not semantically empty (see Jabłońska 2007 on Polish and Dyachkov 2019 on Russian deadjectival verbs). In favour of this claim, I will show that *e*- and *i*-vowels affect the argument structure of the derived verbs and determine their syntactic properties. The basic derivational models involving theme vowels are exhaustively described in Shvedova et al. 1980. For instance, in Russian, the theme vowel *e* derives emission verbs (*bel-e-t'* <white-E-INF> 'be white') or inchoative verbs (*krasn-e-t'* <red-E-INF> 'become white') as well as some predicates denoting transformation (*zver-e-t'* <beast-E-INF> 'become a beast [fig.]'). The theme vowel *i* is used to derive verbs from both nouns and adjectives, and this derivational model is quite productive in modern Russian. Many verbs are causative counterparts to inchoative predicates (*bel-i-t'* <white-I-INF> 'whiten [sth.]')¹. Other classes include some predicates denoting taste that have some stative properties, stative causatives, unergative behaviour-related verbs and verbs derived from names of instruments. All these classes of verbs will be discussed in detail below.

In this paper, I will propose a uniform explanation of the fact that the same theme vowel can give rise to different structural types of verbs. I propose that, semantically, the best way to capture the differences between the two theme vowels in Russian is to implement Ramchand's model (2008) of predicate decomposition. In this model, the verb can be decomposed into up to three subevents {init, proc, res} which are responsible for initial, process and resulting phases. Whereas the result sub-event in Russian is always expressed by perfectivizing lexical prefixes (Svenonius 2004b; Ramchand 2005; Tatevosov 2010), I argue that it is the process sub-event that is encoded by the *e*-vowel and that it is the initial sub-event that is encoded by *i*-vowel. Such an analysis allows to explain many properties of the denominal and deadjectival verbs in question in a uniform way. First, *e*-vowel derives only unaccusative verbs lacking the initial sub-event.

¹ In Russian, causative-inchoative alternation can be also realized by decausativization: *u-glub-i-t'* <PREF-deep-I-INF> 'make sth. deep(er)' – *u-glub-i-t'-s'a* <PREF-deep-I-MED> 'get deep(er)'. I will not consider this derivation in the present paper.

In Ramchand's model, unaccusatives are associated with the [+proc] semantic feature, and I argue that the *e*-vowel can be analysed as the procP head. In contrast, the *i*-theme vowel derives agentive and unergative verbs, and the vowel occupies the position of the initP head.

The paper is structured as follows. First, in § 2, I will describe the types of the verbs derived with *e*- and *i*-vowels. In § 3, I will show how these data can be implemented into the Ramchand's model of predicate decomposition. I argue that the verbs derived with *i*-vowel can appear in two configurations that differ in the number of sub-events they are composed of. I will describe the properties of each semantic type and discuss the question of why *e*- and *i*-vowels cannot co-occur within the same verb form. In § 4, I will discuss the structural types of the verbs derived by adding *i*-vowel to the base. In § 5, I will discuss some more issues concerning the advantages of the present proposal. Specifically, I will compare my proposal to the analysis that was proposed in Jabłońska 2007 for Polish theme vowels. § 6 concludes the paper.

2 Structural Types of Denominal and Deadjectival Verbs

2.1 *e*-Verbs

Verbs derived with *e*-vowel are mostly deadjectival predicates (Shvedova et al. 1980, 344-5). They can be classified as inchoatives (1), but some verbs can be best described as verbs denoting colour emission (Dyachkov 2019, 81-2, 88-9). The latter differ from the former in that they do not encode transition into a state denoted by the adjectival stem. This class is exemplified in (2).

(1)

Jego	lico	bel- e -et.
3SG.POSS	face	white-E-PRES.3SG

'His face is getting white'.

(2)

Na	gorizont-e	par-u	čas-ov	bel- e -l-i	oblak-a.
on	horizon-LOC	couple-ACC	hour-gen.PL	white-E-PST-PL	cloud-PL

'The clouds were whitening on the horizon'.

Apart from deadjectival verbs, in Russian there are several *e*-verbs that are derived from nouns: *satan-e-t'* <satan-E-INF> 'get furious', *zver-e-t'* <beast-E-INF> 'lose control' (Shvedova et al. 1980, 345). However, this class is recognised as non-productive in modern Russian.

2.2 *i*-Verbs

i-vowel derives verbs from both nouns and adjectives. This derivational model is quite productive in modern Russian (Shvedova et al. 1980, 332-5). Many adjectives derive deadjectival verbs, and some of them also derive causative verbs, cf. (3) and (4). However, the theme vowels, as is well-known, do not co-occur within the same verb – that is, causative verbs do not inherit the theme vowel of their inchoative counterparts. Moreover, the event structure of the causative verb does not comprise that of the inchoative one, as I will argue in detail below.

(3)

Tovar	po-dešev- e -l.
merchandise	pref-cheap- E -PST.M

‘The merchandize got cheaper’.

(4)

Prodavec	u-dešev- i -l	tovar.
salesman	pref-cheap- I -PST.M	merchandize.ACC

‘The salesman made the merchandize cheaper’.

Other classes derived with *i*-vowel include some predicates denoting taste that have some stative properties (5),² stative causatives (6),³ unergative behaviour-related verbs (7) and verbs derived from names of instruments (8).

(5)

Jeda	gorč- i -t.
food	sour- I -PRES.3SG

‘The food is sour’.

(6)

Et-o	plat’je	men’a	poln- i -t.
this-N	dress	1SG.ACC	plump- I -PRES.3SG

‘This dress is making me look plump’.

(7)

Čto	ty	tup- i -š?
what	2SG	dumb- I -PRES.2SG

‘Why are you being stupid?’

² Shvedova et al. (1980) do not separate this class from the predicates labeled here as unergatives. I distinguish these two classes because of their significant semantic difference. Below I will show that unergatives are typical eventive verbs, whereas predicates of taste are not.

³ The term “stative causative”, used elsewhere in this paper, refers to causatives that, unlike their dynamic counterparts, do not have any dynamic component and denote state-to-state causal chains, cf. Kratzer 2000; Pyllkkänen 2000; Rothmayr 2009.

- (8)
- | | | |
|-------|--------------------|----------|
| Vas'a | motyž- i -t | zeml'-u. |
| V. | hoe-I-PRES.3SG | soil-ACC |
- 'Vasya hoes the ground'.

The semantic types listed above are the focus of the present paper, although several other types of predicates can be derived using *i*-vowel. Those are exemplified below in (9). Some of these verbs belong to the types that are quite productive, and I will consider them below in § 4.

(9)

	base	denominal/deadjectival verb (infinitive form)	
a.	<i>gost'</i> 'guest'	<i>gost-i-t'</i>	'be a guest (for some time)'
	<i>partizan</i> 'partisan'	<i>partizan-i-t'</i>	'be a partisan'
	<i>bazar</i> 'market, bazaar'	<i>bazar-i-t'</i>	'behave like in a bazaar'
	<i>pudra</i> 'powder'	<i>pudr-i-t'</i>	'powder (verb), cover with powder'
	<i>kaleka</i> 'cripple'	<i>kaleč-i-t'</i>	'make sb. a cripple'
	<i>dym</i> 'smoke'	<i>dym-i-t'</i>	'smoke (verb), fume away'
b.	<i>t'oplyj</i> 'warm'	<i>u-tepl-i-t'</i>	'make sth. warm'
	<i>m'agkij</i> 'soft'	<i>s-m'agč-it'</i>	'make sth. soft'

It turns out that, superficially, the verbs listed in this section do not have any common semantic components. It is obvious that the theme vowel introduces the causative component in the case of deadjectival verbs, as can be seen from (5)-(7), but other verbs like *tup-i-t'* 'behave stupidly' or *gost-i-t'* 'be a guest' do not seem to have any causative semantics. Thus, *i*-vowel cannot be regarded as a pure causative marker. In the following section, I will propose an explanation of the fact that several different semantic types use the same theme vowel, and then I will test predictions made by my proposal.

3 Formal Implementation in Ramchand's Model

For my purposes, I will use the theory of verb decomposition presented in Ramchand (2008). Let us suppose that the initial verb stem can be decomposed into one to three projections – initP, procP and resP – where initP describes the initial state, procP the process phase of the event and resP the result state. The maximal possible structure of a verb is as follows:

(10)



Various structural and semantic types of predicates differ in what projections they include in their nanosyntactic representation. In Ramchand's theory, unaccusatives lack the [+init] feature, since *initP* is associated with the agent's activity and unaccusatives do not have an agent in their semantic representation. If a verb does have *initP*, there are two possibilities. The structure including *initP* but lacking *procP* is interpreted as a stative predicate. The structure having both components is interpreted as an activity. *resP* is responsible for the result state, and only verbs having a result state (that is, achievements and accomplishments) have a *resP* in their representation.

I assume that the structure presented in (10) can be used as the cartographic representation reflecting the (presumably) universal order of syntactic/semantic structure of the predicate. The Ramchandian structure can be mapped onto the traditional *vP/VP* sequence (cf. Jabłońska 2007). There is much debate on the nature of *vP* itself;⁴ many theories assume that the difference in assigning the role of external argument is determined by different 'flavours' of *v*. In such theories, transitive and intransitive verbs are derived independently with different 'flavours' of *v*. For instance, transitive verbs are modelled as structures with v_{TRANS} *P* that project a specifier, and $v_{\text{IN-TR}}$ *P* that do not.

In § 2, I have shown that the class of verbs derived with *i*-vowel is heterogeneous. It comprises both transitive and intransitive predicates, which belong to different semantic and structural types. In order to capture the uniformity of the morphosyntactic makeup of these types (i.e. the fact that these verbs are derived with the same theme vowel), I propose that the theme vowel occupies the same slot in the structure of the verbal form. Thus, the difference between several semantic types arises from a complex interaction of the seman-

⁴ Cf. Marantz 1997; Chomsky 1999; Alexiadou, Anagnostopoulou 2004; Folli, Harley 2007, to mention only a few.

tics of the stem, the theme vowel and the structural configuration of the predicate. Given the fact that the causative/inchoative alternation in Russian is expressed by the change of the theme vowel, I will straightforwardly hypothesise that the theme vowels occupy two different slots in the Ramchandian structure. Specifically, *e*-vowel derives two structural types of verbs, namely emission verbs and inchoative verbs. Both types involve a participant that does not control the action and undergoes the process denoted by the verb. In Ramchand's theory, such predicates cannot have an *initP* and thus have the set of features listed in (11).

(11)

e-vowel: [+proc, ±res]

The crucial generalisation is that these verbs are unaccusative, since the structure represented in (11) corresponds to unaccusative predicates. In Russian, unaccusativity can be revealed by a number of tests (to be discussed below). In contrast, causative verbs alternating with inchoatives do have an *initP* in their representation, and I hypothesise that their structure is as in (12). However, not all the verbs derived with *i*-vowel are causative verbs. Taking into consideration my previous assumption that the uniformity in marking must reflect the structural uniformity, I also hypothesise that there are other configurations where the verbs with *i*-vowel can appear. Let us recall that, for verbs having an *initP*, there is another possible configuration in Ramchand's model. Namely, the *initP* may not have a *procP* complement, and in such cases the verb is interpreted as a stative predicate (13).

(12)

i-vowel: [+init, +proc, ±res]

(13)

i-vowel: [+init]

Thus, the model sketched above makes a crucial prediction. Given the assumption that *i*-vowel is associated with *procP*, all the verbs containing it should not be unaccusative. In the sections below, I will investigate this prediction and explore the semantic properties of the predicates in more detail.

4 Structural Types of *i*-Verbs

4.1 Stative Causatives

Stative causatives in Russian can be derived from a limited number of adjectives. Their list is given in (14) but is possibly not exhaustive. Most of these causatives denote states attributed to a human holder and associated with a visual effect. An example is given in (6) and is repeated below in (15).

(14)

base		deadjectival verb (infinitive form)	
<i>poln-yj</i>	‘plump’	<i>poln-i-t’</i>	‘make sb. look plump’
<i>strojn-yj</i>	‘slim’	<i>strojn-it’</i>	‘make sb. look slim’
<i>star-yj</i>	‘old’	<i>star-i-t’</i>	‘make sb. or sth. look old’

(15)

Et-o	plat’je	men’a	poln-i-t.
this-N	dress	1SG.ACC	plump-I-PRES.3SG

‘This dress is making me look plump’.

Stative causatives have significant properties distinguishing them from true *i*-causatives. First, their core meaning can be described as ‘fake causativization’. For instance, the verb *polnit’* ‘make sb. look plump’ does not denote an event where something makes something plump. In (15), the participant marked by the accusative cannot be described as ‘plump’, because the verb only denotes the visual effect of the dress. This can be seen in the following example, where the property denoted by the verbal stem can be cancelled by the context.

(16)

Ja	xud-aja,	no	et-o	plat’je	men’a	poln-i-t.
1SG	thin-F	but	this-N	dress	1SG.ACC	plump-I-PRES.3SG

‘I am thin, but this dress is making me look plump’.

Second, stative causatives always involve non-animate subjects – for instance, in (15) it is only clothes (but not, say, a human being) that may cause the visual effect. Third, stative causatives usually have eventive counterparts that have all the range of properties of ordinary *i*-causatives. These eventive counterparts are usually prefixed telic verbs denoting enter-into-state encoded by the adjectival base. (17) and (18) are examples of such causatives involving the inanimate and the animate subject, respectively.

(17)

plat'je men'a sil'no u-korot-i-l-o i ras-poln-i-l-o.
 dress 1SG.ACC much PREF-short-I-PST-N and PREF-plump-I-PST-N
 'The dress made me look short and plump'. (Google)

(18)

Vas'a iskusstvenn-o so-star-i-l mebel'.
 V. artificial-ADV PREF-old-I-PST.M furniture.ACC
 'Vasya artificially aged the furniture [lit. made the furniture old]'.

A question arises whether the semantic properties of the stative causatives can be attributed to the configuration of the predicate and its arguments. For instance, if the stative causatives always involve inanimate subjects, can the stative properties be due to the inanimacy itself? The case of (17) suggests that the correct answer is no. Both stative causatives and their telic counterparts involve non-agentive subjects, and in this respect they do not differ from each other. Hence, stativity arises from some other inherent properties of the verbs.

Further investigation shows that stative causatives are predicates denoting actions that can hardly be characterised in terms of durativity – in other words, they do not refer to actions that can be localised within a certain temporal interval. For instance, activities (= predicates having both the *initP* and the *procP*) are non-punctual predicates, and they are compatible with *for*-adverbials, cf. (19). This is not so in the case of stative causatives (20).

(19)

Ja beg-a-l dv-e minut-y / dolgo.
 1SG run-A-PST two-F minute-GEN.PL for.a.long.time
 'I ran for two minutes / *for a long time'.

(20)

* Et-o plat'je men'a poln-i-l-o dv-a dn-ya / dolgo.
 this-N dress 1SG.ACC plump-I-PST-N two-M day-GEN for.a.long.time
 '*This dress was making me look plump for two days / for a long time'.

The example (20) shows that, although stative causatives are atelic predicates, they are not subject to modification by the temporal adverbials that are able to combine with any predicates denoting continuous events. Therefore, they cannot be activities and lack the [+proc] feature. This is consistent with my proposal that statives only have an *initP* in their representation.

Finally, we need to make sure that stative causatives are not unaccusative verbs. For Russian, a number of tests were proposed to

diagnose the unaccusative properties of verbs.⁵ One of the most significant features affecting the unaccusativity of a verb is the animacy of the subject. As shown in Glushan (2009), this feature is responsible for the contrast observed in (21) and (22). In both cases, the verb ‘come’ is used, but it has unaccusative properties and passes the genitive-of-negation test for unaccusativity only in the former case (Pesetsky 1982).

(21)

nikak-ix	pisem	ne	pri-xodi-l-o.
none-GEN.PL	letter.GEN.PL	NEG	PREF-go-PST-N

‘There were no letters coming’.

(22)

??	nikak-ix	l’ud-ej	ne	pri-xodi-l-o.
	none-GEN.PL	people-GEN.PL	NEG	PREF-go-PST-N

‘There were no letters coming’.

Applying some of the unaccusativity tests for Russian, one can see that stative causatives are not unaccusative predicates. The distributive test (Schoorlemmer 2004) shows this. Whereas inchoative verbs derived with *e*-vowel pass this test and can occur in distributive contexts (23), stative causatives cannot (24). This allows us to conclude that the two types of verbs have different argument structures.

(23)

V	každ-oj	korzin-e	za-červiv- e -l-o	po	jablok-u.
in	each-LOC.F	basket-LOC	PREF-wormy-E-PST-N	on	apple-DAT

‘There was an apple in each basket that became worm-eaten’.

(24)

*	Každ-yj	den’	men’a	poln-i-l-o	po	plat’-ju.
	each-M	day	1SG.ACC	plump-I-PST-N	on	dress-DAT

Int.: ‘There was a dress each day that was making me look plump’.

To sum up, I assume that stative causatives do not have a [+proc] feature in their semantic representation and can be regarded as stative predicates. They cannot refer to continuous events and therefore are different from ordinary activities. Simultaneously, they are not unaccusative predicates, which is expected if *i*-vowel is indeed located above the procP.

⁵ Cf. Babyonyshev et al. 2001; Harves 2002; Schoorlemmer 2004; Glushan 2009.

4.2 Predicates of Taste

Let us now turn to the problem of predicates of taste. This class is represented in Russian by the verbs *gorč-i-t* ‘be bitter’ and *kisl-i-t* ‘be sour’. Other basic adjectival stems denoting taste, *solěnyj* ‘salty’ and *sladkij* ‘sweet’, do not derive deadjectival verbs. Although deadjectival verbs derived from adjectives denoting taste do not constitute a productive class, I suggest that the fact that they use *i*-vowel but not another derivational marker fits well in the general picture.⁶

The key question is whether the predicates of taste belong to the class of states or not. If the answer to this question is positive, then they can be modelled as follows:

(25)

Jeda *gorč-i-t*.
food bitter-I-PRES.3SG
‘The food tastes bitter’.



Let us recall that Ramchand (2008) assumes that there are two possible configurations in which the *initP* head appears. If it has a complement (= *procP*), then the structure is interpreted as an eventive predicate. If it lacks a complement, then the structure is interpreted as a stative predicate. I assume that the verbs of taste represent the latter case. The stem is conflated into the *initP* head, and *initP* does not project a *procP* complement.

Like all the verbs having a *procP* but lacking a *resP*, verbs of taste are atelic. They are incompatible with *for*-adverbials (26), the test that can be used to diagnose the *procP* component with the verbs that lack the *resP*.

⁶ An anonymous reviewer points out that the predicates of taste are adjective-like items and are “reminiscent of languages formally having no adjectives”. Given these considerations as well as the fact that these verbs form a minor class, I admit that these verbs could be inherited from an earlier stage of development of the Russian language. Of course, this question needs additional investigation. However, as was shown, the model is not productive and there are no non-derived adjective-like verbs in modern Russian.

(26)

Jed-a	gorč-i-l-a	(*dva	dn'-a).
food-F	bitter-I-PST-F	two	day-GEN

'The food was bitter (*for two days)'.

One can suggest that the incompatibility with *for*-adverbials might be due to other factors than the semantic/syntactic structure. Informally speaking, (26) might be infelicitous, because the state of affairs denoted by the verb does not hold for a limited period of time but is (in some sense) permanent. However, predicates of taste do not pattern with other stative verbs – for instance, with individual level predicates. Individual level predicates are also incompatible with *for*-adverbials (27), but, unlike predicates of taste, they cannot be used in contexts like those exemplified in (28), compared to (29).

(27)

* Vas'a	zna-l	francuzsk-ij	yazyk	dva	god-a.
V.	know-PST.M	French-ACC.M	language.ACC	two	year-GEN

*Vasya knew French for two years'.

(28)

* Vsyak-ij	raz,	kogda	Vasya	vy-ezža-et	
each-M	time	when	V.	PREF-go-PRES.3SG	
za	granic-u,	on	zna-et	francuzsk-ij	yazyk.
behind	border-ACC	3SG	know-PRES.3SG	French-M	language

*Each time Vasya goes abroad, he knows French'.

(29)

Vsyak-ij	raz,	kogda	ty	gotov-iš	jed-u,
each-M	time	when	2SG	prepare-PRES.2SG	food-ACC
on-a	gorč-i-t.				
3SG-F	bitter-I-PRES.3SG				

'Each time you prepare food, it tastes bitter'.

In (29), the episodic interpretation of the predicates of taste is induced by the context. The sentence denotes a situation where, for each occurrence of the event denoted by the main clause, the event denoted by the main clause takes place. Since individual level predicates cannot have episodic interpretations, they cannot be used in such sentences. In contrast, predicates of taste are completely felicitous. This implies that, although predicates of taste denote states, they cannot be treated equally to states denoted by the predicates such as 'know'.

What is the formal way to clear up this discrepancy? The stative causation analysis might be the best way to capture the essential properties of the predicates of taste, which would bring together verbs like *gorčit'* 'be sour' and the stative uses of causatives dis-

cussed in § 4.1. Let us assume that the predicates of taste are stative causatives. Then, in Ramchand's framework, their representation only includes the [+init] feature, and this means that these verbs are not unaccusative. Indeed, this prediction is borne out, since predicates of taste cannot occur in unaccusative contexts, cf. (30).

(30)

*	nikak-oj	jed-y	ne	gorč-i-l-o.
	none- GEN	food-GEN	NEG	bitter- I-PST-N

Int.: 'There was no food which would taste bitter'.

If the predicates of taste are indeed stative causatives, it is natural to assume that individual level predicates, such as 'know', must differ from them in their (nano)syntactic representation. Although I am not aware of any theory that would represent statives of different kinds by different syntactic structures, I hypothesise that individual level predicates have to include more layers of representations than simple statives, since they are more complex semantically. Whereas predicates like 'taste bitter' can have episodic interpretations, individual level predicates can be treated as predicates denoting abstractions over sets of episodic events. However, this question is beyond the scope of this paper and is the matter of a future research.

4.3 Instrumental Verbs

Instrumental verbs with *i*-vowel are represented by a quite large set of items. These items are atelic by default (31) but can be telicized by prefixes (32).

(31)

Vrag	bomb-i-l	gorod	(*za)	dva	čas-a.
enemy	bomb-I-PST	city	in	two	hour-GEN

'The enemy bombed the city (*in) two hours'.

(32)

Vrag	raz-bomb-i-l	gorod.
enemy	PREF-bomb-I-PST.M	city.ACC

'The enemy bombed the city (completely)'.

Thus, instrumental verbs are composed of two sub-events, one of them being the process sub-event and the second the initial sub-event. Thus, the verb is decomposed into procP and initP.⁷

⁷ Such a formal implementation may seem doubtful, since in order to properly derive an instrumental verb, one needs to assume that the stem *bomb-* is first integrated into

There is a strong semantic motivation for the fact that nominal stems denoting instruments can be combined with *i*- but not with *e*-vowel. Taking into consideration the fact that the *e*-vowel lexicalizes predicates lacking initP, I assume that such predicates derived from names of instruments would be semantically impossible. Such predicates would denote actions performed with an instrument but lacking an agent bringing about the whole event, which does not correspond to any possible state of affairs in the real world. Thus, the association of the instrumental verbs with *i*-vowel seems semantically well-motivated.

4.4 Unergative Verbs

A more interesting case is represented by the unergative *i*-verbs. These verbs can be derived both from nouns (33) and from adjectives (34), although the second type is not fully productive and represented by several items in modern Russian.

(33)

base		denominal verb (infinitive form)	
<i>gost'</i>	'guest'	<i>gost-i-t'</i>	'be a guest (for some time)'
<i>rybak</i>	'fisherman'	<i>rybač-i-t'</i>	'be fishing'
<i>batrak</i>	'(farm) labourer'	<i>batrač-i-t'</i>	'work as a farm labourer'

(34)

base		deadjectival verb (infinitive form)	
<i>xitr-yj</i>	'cunning'	<i>xitr-i-t'</i>	'cheat, be cunning'
<i>tup-oj</i>	'stupid'	<i>tup-i-t'</i>	'behave stupidly'
<i>mudr-yj</i>	'wise'	<i>mudr-i-t'</i>	'do something which is too wise'

This class of verbs is not homogenous. Some of the items listed in (33) have properties of activities. For instance, the verb *rybačit'* 'go fishing' is felicitous in atelic contexts describing the agent's activity within a certain period of time (35). In contrast, the activity interpretation of some other items is not available, possibly for pragmatic reasons (36).

a zero procP head and then, for some reasons, moves to initP and combines with *i*-vowel. However, the framework used in this work does not allow to substantiate the movement operations, and I will leave this question open.

(35)

Segodn'a ja rybač-i-l dv-a čas-a.
 today 1SG fisherman-I-PST two-M hour-GEN
 'Today, I was fishing two hours'.

(36)

* Segodn'a on xitr-i-l dv-a čas-a.
 today 3SG cunning-I-PST two-M hour-GEN
 Int.: 'Today, he was being cunning for two hours'.

Another test that can be used to demonstrate that at least some of these verbs belong to the class of activities is their compatibility with the attenuative prefix *po-*. When used with activity verbs, this prefix induces the regular interpretation 'do V for some time'. Crucially, the prefix is compatible with the verbs derived both from nouns and adjectives. This is exemplified below in (37)-(39).

(37)

My nemnogo po-gost-i-l-i u nix.
 1PL a.little ATT-guest-I-PST-PL at 3PL
 'We stayed at their place for a while'.

(38)

My nemnogo po-tup-i-l-i i vz'a-l-i-s' za rabot-u.
 1PL a.little ATT-guest-I-PST-PL and take-PST-PL-MED behind work-ACC
 'We were hanging around (= behaved stupidly) for a while and then got ourselves to work'.

(39)

Molodoj Trezini nemnogo po-xitr-i-l v intervju.
 young T. a.little ATT-cunning-I-PST.M in interview.LOC
 'Young Trezini was being a little cunning in his interview'. (Google)

At the same time, none of the unergative *i*-verbs can appear in unaccusative configurations. All of these verbs include an agentive participant in their representation, which disallows them from appearing in typical unaccusative structures, as shown in § 3. As a consequence, they cannot appear in negation contexts (40).

(40)

?? u nas ni-kogo ne gost-i-l-o.
 at 1PL.GEN NEG-who.GEN NEG guest-I-PST-N
 Int.: 'Nobody stayed at our place'.

These facts allow us to conclude that *i*-verbs can be decomposed into two components, *initP* and *procP*. These verbs are not unaccusative, and there is no evidence that they include only *initP*, like predicates of taste or stative causatives.

The fact that behaviour-related verbs may pattern together with causative verbs was discussed for French by Martin and Piñon (2020). This pattern is shared by Russian unergative verbs, since the same morphological makeup is used to derive both causative deadjectival verbs (see § 3) and unergative predicates. However, there is a significant difference between the Russian and French derivational models. Martin and Piñon show that French unergatives preserve some properties of causative verbs – for instance, they can attach direct objects (41), have anticausative uses etc., which is conditioned by their causative morphosyntax. In contrast, Russian unergative verbs cannot attach direct objects, and the only way to transitivity them is to add a prefix licensing the direct object, cf. (42).

(41)

Sarkozy	diplomat-is-e	le	Hezbollah.
S.	diplomat-VBZ-PRES.3SG	DEF	H.

‘Sarkozy causes Hezbollah to get typical properties of diplomatic organizations’
(Martin, Piñon 2020, 3)

(42)

On	pere-xitr-i-l	/	*xitr-i-l	det-ej.
3SG.M	PREF-cunning-I-PST.M		cunning-I-PST.M	child.PL-ACC

‘He cheated on the children’.

To sum up, unergative verbs in Russian share the morphological makeup with causatives but differ in ability to be used in transitive clauses. Whereas deadjectival causatives are obligatorily transitive, unergatives cannot be transitivized. This property distinguishes them from unergative verbs in some related languages, where these verbs have the causative morphology. However, direct objects can be introduced by some prefixes.

4.5 Relations Between *e*- and *i*-Verbs

In § 3, I showed that the *e*- and *i*-theme vowels do not co-occur within the same verb form. This fact makes some strong predictions, which are in fact borne out by the Russian data. If the structure of the causative verb does not contain that of its inchoative *e*-counterpart, one would expect that the propositions contained in these sentences cannot be reduced to each other. This is true, which can be shown by the following pair, where (43) does not entail (44):

(43)
Vas'a po-bel-i-l sten-u.
V. PREF-white-I-PST.M wall-ACC
'Vasya painted the wall white'.

(44)
Stena po-bel-e-l-a.
wall PREF-white-E-PST-F
'The wall whitened'.

The example (44) is infelicitous because the unaccusative verb *po-bel-e-t'* <PREF-white-E-INF> 'become white(r)' can only be understood as a spontaneous action that cannot have an implicit initiator. Since the 'wall' usually cannot 'whiten' itself, the state of affairs described in (44) seems awkward. In contrast, other uses of the same verb that do not imply any initiator are completely acceptable, cf. (45).

(45)
Jego lico **po**-bel-**e**-l-o ot strax-a.
3SG.POSS face PREF-White-E-PST-N from fear-GEN
'His face whitened from fear in two seconds'.

There is also a significant difference in the semantics of the causative and inchoative counterparts in (43) and (44). Whereas the event described by the verb *pobelet'* 'whiten (intr.)' describes an action of becoming white, the verb *pobelit'* has a more specific sense and usually refers to an action of *painting* something white but not an action of causing the state of affairs, 'the wall whitened'. So, strictly speaking, the two verbs are not direct counterparts to each other.

Moreover, in the case discussed here, the two verbs have the same prefix *po-*, but this is rather an exception than a general rule. In most cases, the inchoative and the causative verb derived from the same adjectival base have different prefixes. This can be illustrated by the following pairs:

(46)
po-dešev-e-t' **u**-dešev-i-t'
PREF-cheap-E-INF PREF-cheap-I-INF
'get cheap(er)' 'make sth. cheap(er)'

(47)
po-star-e-t' **so**-star-i-t'
PREF-old-E-INF PREF-old-I-INF
'get old(er)' 'make sth. old(er)'

(48)

po -tepl- <i>e</i> -t'	u -tepl- <i>i</i> -t'
PREF-warm-E-INF	PREF-warm-I-INF
'get warm(er)'	'make sth. warm(er)'

In each of this pairs, the prefix of the inchoative and the prefix of the causative do not coincide. It is also noteworthy that causative-inchoative alternation is not productive in Russian. Many of the *e*-verbs do not have *i*-counterparts, and vice versa, compare *lys-e-t'* <bald-E-INF> 'get bald' - **lys-i-t*, *tolst-e-t'* <fat-E-INF> 'get fat' - **tolst-i-t'*, *u-lučš-i-t'* <PREF-good.CMPR-I-INF> - **u-lučš-e-t'*, etc.

To sum up, causative and inchoative structural types in Russian cannot be reduced to each other. This is confirmed not only by the fact that *e*-vowel is absent in causatives, but also by semantic tests showing that the structure encoded by causatives is not included into that encoded by inchoatives. We have seen as well that the same (adjectival) base can be merged with different prefixes before they attach a theme vowel. Therefore, predicates of the two types are derived independently of each other.

5 Discussion

Summarising the facts discussed in this paper, I conclude that the verbs derived with *i*-vowel can be subdivided into stative and eventive predicates. The information on the types of verbs that were discussed in this work is given below in (49).

(49)

	stative {init}	eventive {init, proc}
deadjectival causatives		+
stative causatives	+	
predicates of taste	+	
unergative verbs		+
instrumental verbs		+

The problem of syntactic representation of different structural types of verbs is also one of the key topics discussed by Jabłońska (2007), which is the most significant work concerning Slavic theme vowels. Throughout the paper, I modelled the stative verbs as having initP (=vP), while eventive predicates also having procP (=VP). Contrary to this, Jabłońska (2007, 11-12) assumes that the process phase of an event is always located in vP and not in VP. Under her approach, all the dynamic predicates, including unaccusatives, have a preceding state in their semantic representation. Thus, the preceding state (=

$vP = \text{init}P$) has to be postulated not only for causatives, but also for unaccusatives, whereas the latter lack it in Ramchand's model. In order to avoid this mismatch, Jabłońska claims that vP denotes a process or a state. Thus, two different "flavours" of vP , namely vP_{stative} and vP_{process} , are postulated.

Jabłońska shows that some Polish object experiencer verbs (*dziw-i-ć* 'surprise', *złość-i-ć* 'irritate') are also derived with *i*-vowel. She argues that those can be represented as in (50). This structure is identical to the structure of causative verbs (51), with the exception that vP represents a state rather than a process.

(50)

$[vP_{\text{stative}}, VP_{\text{become}}, RP]$

(51)

$[vP_{\text{process}}, VP_{\text{become}}, RP]$

However, there are some challenges facing the parallelism of the two structures. Indeed, Polish object experiencer verbs encode enter-into-states, but the cases of stative causatives and predicates of taste seem more complicated. Although Jabłońska does not consider in detail the properties of experiencer verbs, it seems that they are subject to the same processes as causative verbs. For instance, the fact that the experiencer is the holder of the result state can be confirmed by some language-specific tests. For instance, any holder of the state in Slavic languages can appear as the subject of the passive construction (cf. English *his behaviour surprised Mary* vs. *Mary was surprised by his behaviour*). This can be seen in (52) and (53) presenting the Russian experiencer verb *udivit'* 'surprise', where (53) is the passive counterpart of (52).

(52)

Jego	povedenije	u-div-i-l-o	men'a.
3SG.POSS	behaviour	PREF-surprise-I-PST-N	1SG.ACC

'His behaviour surprised me'.

(53)

Ja	by-l	u-divl-ën	jego	povedeni-jem.
1SG	be-PST	PREF-surprise-PTCP.PASS	3SG.POSS	behaviour-INST

'I was surprised by his behaviour'.

The situation is different with predicates of taste and stative causatives. The former are intransitive predicates and thus cannot be passivized. Stative causatives are transitive (54), but transforming them into passives results in an ungrammatical structure (55).

(54)

plat'je	men'a	ras-poln-i-l-o.
dress	1SG.ACC	PREF-plump-I-PST-N

'The dress made me look short and plump'.

(55)

*	Ja	by-l-a	ras-poln-en-a	plat'-jem.
	1SG	be-PST-F	PREF-plump-PTCP.PASS-F	dress-INST

'*I was made look plump by the dress'.

However, with some stative causatives, passivization is possible. This is the case of the verb *starit'* 'make smth. old'. It can be used as a stative causative, but its prefixed counterpart can appear in passive constructions (56). Nevertheless, there is a subtle semantic difference between the two verbs – the prefixed verb cannot be used as a counterpart to the stative causative. Therefore, it cannot have the meaning 'make sb. look old' (57).

(56)

Mebel'	by-l-a	iskusstvenn-o	so-star-en-a.
furniture	be-PST-F	artificial-ADV	PREF-old-PTCP.PASS-F

'The furniture was artificially aged'.

(57)

*	Ja	by-l-a	so-star-en-a	plat'-jem.
	1SG	be-PST-F	PREF-old-PTCP.PASS-F	dress-INST

'*I was made look old by the dress'.

If the passivization test is indeed the test sensitive to the presence/absence of the result state in the structure of the predicate, then the event structure of stative predicates needs to be reconsidered. At the same time, Polish experiencer verbs and stative causatives explored in this paper belong to classes with different semantic properties, and no direct parallels can be drawn between them. However, if stative causatives constitute a separate class on their own, then their properties must be properly integrated into the model. I suggest that the assumption that they only have an *initP* in their structure is the best way to do this. This structure implies that stative causatives do not have a result state. Apart from the passivization test, there are other empirical arguments in favour of this claim. Specifically, stative causatives do not assign the property denoted by the adjectival base to the subject – that is, if a dress makes somebody look plump, it does not mean that somebody *is* plump, cf. (16). Consequently, there is even less evidence that such predicates encode result states.

Alternatively, Jabłońska (2007) claims that theme vowels can be subdivided into “high” and “low” themes. Low themes spell out no more than up until VP_{Become} projection and correspond to the unaccusative type, whereas high themes lexicalise the vP-VP_{Become} functional sequence. In other words, *i*-vowel encodes both the process phase and the BECOME component (“transition” in Jabłońska’s terms). The representations for both types are given in (58) and (59).

(58)

e-vowel: [+VP]

(59)

i-vowel: [+vP, +VP]

This proposal would allow us to account for the fact that the theme vowel merges directly with the verbal root and escape the necessity to postulate movement operations for instrumental verbs. In Ramchand’s model there is no formal mechanism that would allow to ‘skip’ the procP component and to attach the adjectival stem (possibly combined with the prefix) directly to the theme vowel. However, Jabłońska makes an additional assumption that in unergative structures, *i*-vowel does not include the VP level in its representation. This assumption may resolve the problem of predicates of taste – if they are regarded as unergative verbs, they do not include the VP level in their structure. The drawback of this proposal is that stative causatives are not unergative, and thus they have to include the VP level (= transition phase) in their structure, which is not consistent with the facts discussed in this section.

The last question that remains open both in Jabłońska’s and in my proposal is the way the stative vs. eventive interpretation is induced. We have seen that some verbs (namely, predicates of taste) are always stative, whereas some causatives can be both stative and eventive. Although these facts can be formalised via feature underspecification ({+init} vs. {+init, +proc}), the exact mechanism of this process is yet to be studied and remains beyond the scope of this paper.

6 Conclusion

In this paper, I have considered the properties of two Russian theme vowels that are used to derive denominal and deadjectival verbs. I have shown that there are structural factors determining the syntactic properties of these predicates. In particular, *e*-vowel derives only unaccusative predicates, whereas *i*-vowel derives predicates that can only be non-unaccusative. I explored the properties of several types of *i*-verbs, namely true causatives, stative causatives expressing fake

causativization, unergatives, predicates denoting taste properties, and denominal instrumental verbs, and showed that they can be divided into two major subtypes. Some of these predicates are evenitive, whereas others are stative. I assumed that this is due to the fact that the theme vowel occupies the same position in the structure of a verbal form but appears in two different syntactic configurations, which can be implemented into Ramchand's theory of predicate decomposition as {+init} and {+init, +proc} structures.

Abbreviations

1, 2, 3	1 st , 2 nd , 3 rd person
a, e, i	theme vowels
ACC	accusative
ADV	adverbializer
ATT	attenuative
CMPR	comparative
AT	dative
DEF	definite
F	female
GEN	genitive
INF	infinitive
INST	instrumental
LOC	locative
M	male
MED	middle voice
N	neuter
NEG	negation
PASS	passive
PL	plural
POSS	possessive
PREF	(telicizing) prefix
PRES	present
PST	past
PTCP	participle
SG	singular

Bibliography

- Alexiadou, A.; Anagnostopoulou, E. (2004). "Voice Morphology in the Causative-inchoative Alternation: Evidence for a Non-unified Structural Analysis of Unaccusatives". Alexiadou, A.; Anagnostopoulou, E.; Everaert, M. (eds), *The Unaccusativity Puzzle. Explorations of the Syntax-Lexicon Interface*, 114-36. Oxford: Oxford University Press. Oxford Studies in Theoretical Linguistics. <https://doi.org/10.1093/acprof:oso/9780199257652.003.0005>.
- Babyonyshev, M. et al. (2001). "The Maturation of Grammatical Principles: Evidence from Russian Unaccusatives". *Linguistic Inquiry*, 32(1), 1-44. <https://doi.org/10.1162/002438901554577>.
- Borer, H. (2005). *Structuring Sense. An Exo-Skeletal Triloggy*. Oxford: Oxford University Press.
- Chomsky, N. (1999). "Derivation by Phase". Kenstowicz, M. (ed.), *Ken Hale. A Life in Language*. Cambridge: MIT Press, 1-52.
- Dyachkov, V. (2019). *Tipologija deadjektivnyx glagolov* (Typology of Deadjectival Verbs) [PhD dissertation]. Moscow: Institute of Linguistics, RAS.
- Glushan, Z. (2009). *On Animacy and Unaccusativity in Russian*. Paper Presented at the 8th Conference on Formal Description of Slavic Languages (University of Potsdam, 2-5 September).
- Harves, S. (2002). *Unaccusative Syntax in Russian* [PhD dissertation]. Princeton: Princeton University.
- Folli, R.; Harley, H. (2007). "Causation, Obligation and Argument Structure: On the Nature of Little *v*". *Linguistic Inquiry*, 38, 197-238. <https://doi.org/10.1162/ling.2007.38.2.197>.
- Jabłońska, P. (2007). *Radical Decomposition and Argument Structure* [PhD dissertation]. Tromsø: University of Tromsø.
- Kratzer, A. (2000). "Building Statives". Conathan, L.J.; Good, J.; Kavitskaya, D.; Wulf, A.B.; Yu, A.C.L., *Proceedings of the 26th Meeting of the Berkeley Linguistics Society*. Berkeley (CA): Berkeley Linguistics Society, 385-99.
- Lunt, H.G. [1955] (2001). *Old Church Slavonic Grammar*. 7th ed. Berlin; New York: Walter de Gruyter.
- Marantz, A. (1997). "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon". *Penn Working Papers in Linguistics*, 4, 201-25.
- MacDonald, J.E. (2006). *The Syntax of Inner Aspect* [PhD dissertation]. New York: Stony Brook University.
- Martin, F.; Piñón, C. (2020). "Verbalizing Nouns and Adjectives: The Case of Behavior-related Verbs". *Glossa: A Journal of General Linguistics*, 5(1). <https://doi.org/10.5334/gjgl.833>.
- Pesetsky, D. (1982). *Paths and Categories* [PhD dissertation]. Cambridge: MIT.
- Pylkkänen, L. (2000). "On Stativity and Causation". Tenny, C.; Pustejovsky J. (eds), *Events as Grammatical Objects: The Converging Perspectives of Lexical Semantics and Syntax*. CSLI Publications, 417-45.
- Ramchand, G. (2004). "Time and the Event: The Semantics of Russian Prefixes". *Nordlyd*, 32, 323-61.
- Ramchand, G. (2008). *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486319>.
- Rothmayr, A. (2009). *The Structure of Stative Verbs*. Amsterdam: John Benjamins.

- Schoorlemmer, M. (2004). "Syntactic Unaccusativity in Russian". Alexiadou, A.; Anagnostopoulou, E.; Everaert, M. (eds), *The Unaccusativity Puzzle. Explorations of the Syntax-Lexicon Interface*. Oxford: Oxford University Press, 114-36. Oxford Studies in Theoretical Linguistics. <https://doi.org/10.1093/acprof:oso/9780199257652.003.0009>.
- Shvedova, N. et al. (1980). *Russkaja grammatika*, Tom 1 (A Russian Grammar. Volume 1). Moscow: Nauka.
- Svenonius, P. (2004a). "Slavic Prefixes and Morphology: An Introduction to the Nordlyd Volume". Svenonius, P. (ed.), *Nordlyd 32 (2): Special Issue on Slavic Prefixes*. Tromsø: University of Tromsø, 177-204.
- Svenonius, P. (2004b). "Slavic Prefixes Inside and Outside VP". Svenonius, P. (ed.), *Nordlyd 32 (2): Special Issue on Slavic Prefixes*. Tromsø: University of Tromsø, 205-53.
- Tatevosov, S. Татевосов, С. (2010). *Aktsional'nost' v leksike i grammatike* Акциональность в лексике и грамматике (Aktionsart in Lexicon and Grammar) [PhD dissertation]. Moscow: Moscow State University.
- Vaillant, A. (1966). *Grammaire comparée des langues slaves*. Tome III, *Le verbe*. Paris: Klincksieck.

Latino o cirillico? Le proposte di un alfabeto misto nelle visioni di unità jugoslava del periodo interbellico

Giustina Selvelli

Università Ca' Foscari Venezia, Italia

Abstract This paper discusses the context of script choice (Latin and Cyrillic) in the Kingdom of Yugoslavia in the late 1920s and early 1930s, focusing on bialphabetism and biliteracy as official policies of the country. I place the topic in the framework of Latinization in the inter-war period and examine three texts by Yugoslav authors that propose a 'hybrid' writing system containing the characters of both alphabets as a solution to digraphia. It then explores some reactions to such proposals, including the one of Aleksander Belić. The article is based on the analysis of previously unknown sources found in the Matica Srpska Library in Novi Sad.

Keywords Digraphia. Yugoslav alphabet. Cyrillic script. Latin script. Kingdom of Yugoslavia.

Sommario 1 Introduzione: il serbo-croato in due alfabeti: digrafia, bialfabetismo e bialfabetizzazione. – 2 Il contesto di latinizzazione del primo dopoguerra. – 3 Le prime proposte di 'sistema misto' apparse sulla rivista *Život i rad*. – 4 L'alfabeto jugoslavo di Pavle Ž. Radivojević. – 5 Alcune reazioni alle proposte della nuova scrittura mista: Arandelović e Belić. – 6 La reazione di Danilo A. Živaljević e l'applicazione pratica delle proposte di scrittura. – 7 Valutazione critica. – 8 Conclusioni.



Peer review

Submitted	2021-11-02
Accepted	2021-12-09
Published	2020-12-21

Open access

© 2021 | Creative Commons Attribution 4.0 International Public License



Citation Selvelli, S. (2021). "Latino o cirillico? Le proposte di un alfabeto misto nelle visioni di unità jugoslava del periodo interbellico". *Balkanica et Slavia*, 1(2), [1-20] 205-224.

DOI 10.30687/BES/0/2021/02/003

1 Introduzione: il serbo-croato in due alfabeti: digrafia, bialfabetismo e bialfabetizzazione

All'interno del Regno dei Serbi, Croati e Sloveni, poi Regno di Jugoslavia (1918-41), così come nella Jugoslavia socialista, i fattori 'scismatici' più salienti a livello linguistico erano incarnati dalle varianti ekava/ijekava della lingua serbo-croata così come dalla compresenza dell'alfabeto latino e cirillico. Entrambi questi aspetti giocarono un ruolo significativo nei dibattiti tra quelle che potrebbero essere descritte come due fazioni opposte: quella più nazionalista, che impiegava questi elementi per enfatizzare la diversità e l'incompatibilità fra serbi e croati, e quella 'pro-unitaria' (linguisticamente, politicamente e culturalmente), che spesso si scontrava con tali elementi della lingua, percependoli come un ostacolo alla realizzazione dei suoi ideali unificanti. In una certa misura, queste divisioni e antagonismi permasero attivi fino alla fine della Jugoslavia (e oltre).

In seguito alla fine della Prima guerra mondiale, nel contesto sociopolitico successivo alla creazione del Regno dei Serbi, Croati e Sloveni, la questione dell'alfabeto cominciò a porsi per la prima volta a livello istituzionale nella nuova entità statale comune. Il risultato fu la formalizzazione di una condizione di digrafia sincronica (*synchronic digraphia*, si veda Dale 1980, 5)¹ che era già stata definita nella Dichiarazione di Corfù del 1917. Come primo passo ufficiale verso l'istituzione dello Stato comune, la dichiarazione affermava infatti che l'alfabeto cirillico e quello latino erano uguali davanti alla legge (cf. Krfska Deklaracija 1917).²

Tuttavia, tale riconoscimento non bastò a placare i dibattiti relativi alle questioni di scrittura; in particolare, si discusse molto sulla situazione di digrafia o 'bialfabetismo', definita con il termine di *dvoazbučnost* (Bugarski 1997, 38, 91) in serbo-croato, così come sulla situazione di 'bialfabetizzazione' ('biliteracy' in inglese, si veda Bunčić 2016, 68), termine che qui utilizzo per descrivere la situazione in cui entrambi gli alfabeti sono usati in modo intercambiabile da una o più parti. Questa era infatti la situazione originariamente prevista nel nuovo Regno, che poi persistette in una certa misura nella Seconda Jugoslavia, e che è tuttora attuale nella Repubblica di Serbia.

Nel 1929, il regime del re Aleksandar (la cosiddetta *Šestojanuarska diktatura*, stabilitasi in quell'anno) confermò per legge l'eguale sta-

¹ In tempi più recenti, oltre al termine *digraphia*, i termini inglesi *biscriptality* (Bunčić 2016) e *biscriptalism* (si vedano Greenberg 2004, 41; Feldman, Barac-Cikoja 1996) sono emersi per descrivere lo stesso fenomeno, e sono oggi molto più utilizzati.

² Si veda il 6° principio: «Entrambi gli alfabeti, cirillico e latino, sono completamente uguali e ognuno è libero di usarli in tutto il Regno. Tutte le autorità statali e di autogoverno hanno l'obbligo e il diritto di utilizzare entrambi gli alfabeti, secondo la volontà dei cittadini» (Krfska deklaracija, *Srpske Novine*, 83, Beograd, 13 luglio 1917).

tus di latino e cirillico attraverso la pubblicazione delle *Istruzioni ortografiche per tutte le scuole primarie, secondarie e professionali nel Regno dei Serbi, Croati e Sloveni*,³ che prevedevano l'obbligo di imparare entrambi i sistemi di scrittura a partire dalla terza classe della scuola primaria (Pravopisno upustvo 1929), sostenendo così il bialfabetismo come pratica prioritaria. Tuttavia, l'imposizione di tale obbligatorietà non era vista da tutti in maniera favorevole; infatti, oltre all'idea di bialfabetizzazione, tre principali ideologie di scrittura stavano prendendo forma in quegli anni.

La prima era sostenuta da coloro che consideravano il cirillico un impedimento al raggiungimento di una vera unificazione nazionale, in linea con le tendenze ideologiche di latinizzazione attive negli altri stati balcanici (Selvelli 2018), così come nel più ampio spazio eurasiatico.

Nel 1914, infatti, Jovan Skerlić, eminente critico letterario serbo, espresse la volontà di trovare una soluzione compromesso che avvicinasse maggiormente serbi e croati sul piano linguistico, incoraggiando un famoso dibattito che coinvolse diversi intellettuali (Banac 1984, 211). Secondo Skerlić, era opportuno che i serbi abbandonassero l'alfabeto cirillico nel contesto di utilizzo pubblico, e che i croati adottassero la variante *ekava* della pronuncia al posto di quella *ijekava*: in tal modo il serbo-croato sarebbe diventato una lingua propriamente unificata (Skerlić 1914). Tuttavia, sia la morte prematura di Skerlić nel 1914 che lo scoppio della Prima guerra mondiale di poco successiva interruppero il dibattito, che venne ripreso solo dopo la fine del conflitto.

A quanto pare, il cirillico era considerato un problema anche dallo stesso Re Aleksandar, il quale aveva previsto l'abolizione di questo sistema di scrittura nell'interesse dello jugoslavismo⁴ (Stefanović 2015, 112), un piano che tuttavia non venne mai implementato.

La seconda ideologia di scrittura corrispondeva all'opinione di studiosi (prevalentemente di origine serba) di introdurre l'alfabeto cirillico come sistema di scrittura esclusivo in tutti i territori del Regno di Jugoslavia, sottolineando il suo valore per la storia culturale slava, e legittimandone l'autorevolezza anche attraverso l'opera di riforma ortografica realizzata dal filologo Vuk Stefanović Karadžić un centinaio di anni prima.

In aggiunta a queste principali correnti, tra la fine degli anni Venti e l'inizio degli anni Trenta, si sviluppò un'ulteriore ideologia di scrittura che, seppure marginale, si presentava come molto singolare. Questa era sostenuta da coloro che affermavano la necessità di svi-

³ Il *Pravopisno upustvo za sve osnovne, srednje i stručne škole u kraljevini SHS*.

⁴ A quanto pare, l'opinione dell'eminente linguista Aleksandar Belić aveva fatto desistere il re dal realizzare tale progetto (si veda Stefanović 2015, 5).

luppare un nuovo alfabeto costituito da una combinazione di caratteri dell'alfabeto cirillico e latino. Partiremo ora dalla descrizione del contesto ideologico di latinizzazione per poi dedicare uno spazio speciale all'ultimo filone ideologico, quello a supporto del cosiddetto 'alfabeto jugoslavo'.

2 Il contesto di latinizzazione del primo dopoguerra

Come attestazione dell'influenza degli ideali di unificazione alfabetica a livello globale nel periodo interbellico, è essenziale menzionare l'importante lavoro della Commissione Internazionale per la Cooperazione Intellettuale. Nel 1929, questo organismo della Società delle Nazioni decise di intraprendere uno studio scientifico sulla possibilità di promuovere l'uso dei caratteri latini in tutto il mondo per poter pervenire a una migliore comprensione e comunicazione reciproche tra i vari stati, soprattutto fra il mondo occidentale e l'Asia (Société des Nations 1934). La Commissione inaugurò le sue attività in un momento in cui in molti paesi dell'ampio spazio eurasiatico si affermava un diffuso ottimismo su un possibile cambiamento in ambito grafico in questa direzione, incoraggiato dai casi di successo di riforme dell'alfabeto come quella turca e da quelle che erano in corso presso vari popoli dell'Unione Sovietica, così come dallo sviluppo di schemi di trascrizione in caratteri latini per le lingue cinese e giapponese. Cinque anni dopo, nel 1934, la Commissione pubblicò il rapporto sullo stato della latinizzazione in diversi paesi del mondo intitolato *L'adoption universelle des caractères latins*, contenente esempi di riforme già realizzate oppure considerate imminenti o possibili. Nella sua introduzione al volume, il linguista danese Otto Jespersen sosteneva come, malgrado le sue imperfezioni e difetti, l'alfabeto latino fosse l'unico di cui si potesse raccomandare l'adozione universale, in virtù della maggiore chiarezza dei suoi caratteri nelle pratiche di lettura e scrittura e la sua maggiore idoneità alla stampa rispetto ad altri sistemi di scrittura. La ragione più decisiva per la sua 'universalizzazione', tuttavia, era il fatto che l'uso di tale alfabeto risultava essere una pratica saldamente consolidata in Occidente, nei paesi «più importanti per l'intera civiltà mondiale»⁵ (Jespersen 1934, 13). Jespersen sosteneva tale riforma sulla base di una sorta di idealismo, affermando inoltre il seguente:

Non vi è dubbio che la cooperazione intellettuale in tutto il mondo civilizzato verrebbe estremamente facilitata se lo stesso sistema di scrittura venisse impiegato ovunque; la varietà di alfabeti

⁵ Una visione che oggi appare pregiudicata da un alto grado di eurocentrismo.

in uso è infatti uno dei maggiori ostacoli alla riconciliazione tra le nazioni e le razze.⁶ (13)

Il linguista denunciava inoltre le forze che si opponevano a tale riforma alfabetica nei paesi che possedevano un sistema di scrittura diverso da quello su base latina: queste si nutrivano non solo di un conservatorismo «fortemente radicato nella natura umana» ma anche del nazionalismo che rifiutava di adottare un alfabeto «preso in prestito» da un'altra nazione (13). In tale contesto, il linguista citava il caso della Jugoslavia, dove si poteva osservare lo 'strano spettacolo' di una popolazione scissa in due sfere religiose facenti uso di due sistemi di scrittura distinti per trascrivere quella che «in realtà era la stessa lingua» (13). Jespersen traeva paragoni fra il caso jugoslavo e quello indiano, dove la lingua indostana era divisa nella sua forma maomettana, cioè l'urdu, e quella brahmanica, l'hindi, utilizzando due alfabeti distinti: quello su base araba e quello su base devanagari.⁷

Nello stesso volume, alla sezione dei rapporti sullo stato della latinizzazione in alcuni paesi europei e asiatici, troviamo anche un breve testo sulla Jugoslavia, estratto da una lettera del linguista e slavista francese André Vaillant, datata 14 aprile 1934. Qui, lo studioso sosteneva che «l'esistenza dei due alfabeti in Jugoslavia è impraticabile e tutti ne soffrono gli effetti, soprattutto perché questi due alfabeti sono quasi identici nella forma di alcune lettere» (Vaillant 1934, 184). Lo studioso ricordava anche i tentativi compiuti in Serbia prima della guerra per rimediare a questa situazione (riferendosi, anche se non esplicitamente, a Jovan Skerlić), attraverso la proposta di adottare la *latinica* croata come sistema di scrittura esclusivo. A questo proposito, tuttavia, Vaillant rimarcava come l'opinione pubblica serba non sembrasse affatto pronta ad abbandonare l'alfabeto cirillico. Ciò appare infatti evidente da una serie di articoli pubblicati in alcuni influenti giornali in quegli anni, tra cui il *Letopis Matice Srpske* di Novi Sad, che testimoniano gli sforzi paralleli di importanti intellettuali serbi per difendere l'alfabeto cirillico (si veda ad esempio Stošić 1931). La questione si era leggermente sviluppata

⁶ Se non diversamente specificato, tutte le traduzioni sono dell'Autrice.

⁷ Collin (2011, 40) ha affermato come la condizione di digrafia si verifichi generalmente nei contesti in cui una comunità linguistica appare divisa tra due diverse identità religiose o tra due identità nazionali in conflitto fra di loro: secondo l'autore, si tratterebbe quindi di un fenomeno foriero di profonde conseguenze politiche. Lo studioso R.D. King ha inoltre affermato che «la digrafia è generalmente un segno esteriore e visibile dell'odio etnico o religioso» (2001, 44), considerazione che appare decisamente controversa, dal momento che, nel corso dei secoli, sui territori balcanici sono coesistiti non due, ma addirittura quattro o più diversi sistemi di scrittura (se includiamo il glagolitico, la *arebica* e l'alfabeto *bosančica*, o *hrvatska ćirilica*) per trascrivere varianti della stessa lingua, senza che ciò generasse problemi di tolleranza a livello etnico o religioso.

in tale direzione «in virtù dei contatti più stretti tra serbi e jugoslavi nelle province occidentali» (Vaillant 1934, 184). Lo studioso francese ricordava inoltre la propensione del re Aleksandar a imporre un unico alfabeto, quello latino:

A un certo punto la dittatura aveva annunciato un progetto di unificazione dei due alfabeti, chiaramente a vantaggio dell'alfabeto latino, ma sembra che si sia ritirato di fronte alle difficoltà nell'implementazione di questo progetto così radicale. (184)

Il problema, insomma, risultava essere di natura sia politica che religiosa: si trattava certamente di una sfida complessa.

3 Le prima proposte di 'sistema misto' apparse sulla rivista *Život i rad*

La situazione descritta dal linguista Otto Jespersen in riferimento alla lingua serbo-croata era rappresentativa della condizione di 'digrafia sincronica' alla quale lo studioso vedeva rimedio solo nell'adozione esclusiva dell'alfabeto latino. Come notato sopra, la politica alfabetica adottata nel Regno dei Serbi, Croati e Sloveni, e più tardi nel Regno di Jugoslavia (con l'eccezione dell'idea di breve durata del re Aleksandar di eliminare il cirillico), sembrava invece molto più mirata a creare una forma di 'bifabetizzazione' statale, attraverso la quale ciascun cittadino sarebbe stato in grado di scrivere e leggere entrambi gli alfabeti.⁸

In quegli anni, tuttavia, vi fu anche chi, nella visione di una politica di alfabetizzazione unificante a supporto del nuovo stato comune, si spinse al punto di elaborare una particolare proposta di 'sistema misto', fatto rarissimo se non unico nella storia dei sistemi di scrittura. Armando Petrucci ha affermato che, a differenza delle lingue,

i sistemi grafici oggi in uso appaiono fortemente impermeabili gli uni agli altri, spesso anche, se non soprattutto per ragioni ideologiche e politiche di prestigio e di identificazione nazionale. (Petrucci 2007, 52)

Il caso del cosiddetto alfabeto jugoslavo, sebbene mai messo in pratica ma solo teorizzato, sembra minare fundamentalmente questo

⁸ Un fatto anche dimostrato dalla pubblicazione di diverse opere sul tema dell'apprendimento di entrambi gli alfabeti che apparvero in quegli anni, tra cui: *Način učenja latinice posle ćirilice i ćirilice posle latinice*, di Jovan Udicki, pubblicato a Novi Sad nel 1921, e *Album šrifta: ćirilica i latinica*, di T.K. Pogoreljski Perehud, pubblicato a Zemun nel 1924.

principio basilare del funzionamento dei sistemi di scrittura in epoca moderna. Può essere così definito come un caso di ‘script-mixing’ o ‘grapho-hybridization’ (Rivlina 2016), ovvero un ‘sistema grafico ibrido’, un sistema misto, ottenuto «attraverso il trasferimento di caratteri non da un unico sistema di scrittura, ma da più sistemi differenti» (Baglioni, Tribulato 2015, 22).

Uno dei principali promotori della soluzione di un ‘alfabeto misto’ nel periodo tra le due guerre fu lo studioso serbo Božidar Stojanović, il quale nell’autunno del 1934 pubblicò sulla rivista *Život i rad*⁹ un articolo intitolato «Jugoslovenska azbuka» (L’alfabeto jugoslavo). Le ragioni per cui Stojanović difendeva l’introduzione di un nuovo sistema di scrittura erano numerose, ovviamente tutte di natura ideologica e in un certo senso utopica, legate alla facilitazione dell’unificazione tra serbi, croati e sloveni, che a suo avviso costituivano un’unica nazione per «sangue e lingua» (Stojanović 1934, 724). L’autore notava come la coesistenza di due alfabeti costituisse il maggiore ostacolo al raggiungimento degli obiettivi unificanti; tuttavia, tale situazione poteva essere superata attraverso dei «sacrifici accettabili». Due alfabeti costituivano un «lusso», e contribuivano allo sviluppo di forme di antagonismi etnici; inoltre, confondevano i bambini, rendendo la loro educazione, così come quella delle minoranze nazionali, ben più complicata (726).

Per quanto riguarda il cirillico, l’autore ne riconosceva l’importanza per il popolo serbo, notando come in esso fossero presenti «sia attributi tribali che tradizioni religiose»: questo, sosteneva lo studioso, poteva essere definito come «alfabeto serbo-ortodosso», ed era pertanto impensabile renderlo l’unico sistema di scrittura in Jugoslavia. A dispetto del suo uso diffuso in tutto il mondo, tuttavia, lo stesso valeva per l’alfabeto latino, poiché incarnava un elemento esplicitamente «croato» e «cattolico» nel paese (727).

Dal momento che qualsiasi tentativo di introdurre solo l’uno o l’altro alfabeto avrebbe provocato inevitabili attriti e sarebbe stato destinato al fallimento, Stojanović raccomandava l’adozione di un nuovo sistema di scrittura per trascrivere la lingua serbo-croata, qui definita come ‘lingua jugoslava’. L’alfabeto jugoslavo era estremamente facile da imparare e consisteva in una combinazione di caratteri cirillici e latini che «garantiva il principio di equità fino alla precisione matematica», in modo da non favorire un alfabeto a scapito dell’altro (730). Metà delle lettere sarebbero state tratte da un alfabeto e la restante metà dall’altro, in aggiunta ai caratteri condivisi. Stojanović specificava così alcuni dettagli pratici e tecnici dell’introduzione di questo nuovo sistema di scrittura nel paese:

⁹ Rivista che si occupava di temi ‘socio-letterari’. Non è chiaro il motivo per cui questa rivista sia stata scelta come sede per il dibattito sull’alfabeto jugoslavo.

Si deve ordinare che tutte le stamperie e le macchine da scrivere esistenti vengano riorganizzate e adattate alle esigenze dell'alfabeto jugoslavo. I libri e i testi scolastici, tutti i periodici, i giornali, l'amministrazione statale e in generale tutto ciò che è di carattere pubblico devono essere stampati e scritti esclusivamente nell'alfabeto jugoslavo. (732)¹⁰

L'introduzione del nuovo alfabeto avrebbe implicato numerosi vantaggi a livello pratico di alfabetizzazione: ad esempio, se il cittadino medio jugoslavo avesse desiderato accedere alle opere classiche non ancora pubblicate nella nuova edizione jugoslava, questo non sarebbe risultato troppo difficile, poiché sarebbe stato sufficiente imparare le restanti dodici lettere del cirillico o del latino, con uno sforzo minimo. Infatti, attraverso l'alfabetizzazione nel nuovo sistema di scrittura jugoslavo, il popolo avrebbe automaticamente acquisito familiarità con «tre quinti» di entrambi i sistemi di scrittura. L'alfabeto jugoslavo di Stojanović soddisfaceva impeccabilmente il principio fonemico, presupponendo l'eliminazione di tutti i digrammi presenti nella *latinica* croata. Esso si componeva delle seguenti lettere:

a	б	в	g	d	ђ	e	ž	z	i	j	k	л	љ	m
n	њ	о	п	р	s	t	ћ	u	f	x	с	ч	ц	š

Mettendo a confronto il contesto locale di scelte alfabetiche con quello dell'abbandono dell'alfabeto arabo in Turchia, Stojanović lodava in particolare i risultati ottenuti da Atatürk, essendo egli riuscito a sostituire il «difficile», «superato» e «scomodo» sistema di scrittura arabo con un alfabeto straniero nuovo per le masse ma ben più perfetto. Egli commentava:

E quando quel tale popolo abbandona tutto ciò che si pone come ostacolo alla via del progresso e di un futuro migliore, adottando addirittura un alfabeto del tutto sconosciuto e straniero, che ha dovuto imparare per mesi, [...] perché dovremmo noi inciampare da fermi, perché non dovremmo sostituire i nostri due alfabeti con uno jugoslavo, privo di qualsiasi lettera straniera, dal momento che è di enorme rilevanza e interesse per l'intero nostro popolo, e incomparabilmente più facile e accettabile? (733)¹¹

10 «Треба наредити, да се све штампарије и постојеће писаће машине преуреде и оспособе за употребу Југословенске азбуке. Морале би се књиге и уџбеници, сви часописи, листови, државна администрација и у опште све што има карактер јавности штампати и писати само Југословенском азбуком».

11 «Па кад тај такав народ напушта све што му стоји као препрека на путу прогреса и боље будућности, и прихвата чак и туђу сасвим непознату азбуку, коју

Eventuali obiezioni a tale riforma, sosteneva sempre Stojanović, non potevano inficiare i vantaggi fondamentali derivanti dall'introduzione dell'alfabeto jugoslavo, poiché questo avrebbe portato non solo all'affermazione di una lingua comune, ma anche di una cultura comune e dello jugoslavismo in generale (731). Per quanto riguarda l'impatto economico, la sua introduzione avrebbe implicato il minor sacrificio possibile per i popoli coinvolti, in contrasto con la situazione risultante dall'uso simultaneo di entrambi gli alfabeti, da cui derivavano enormi danni per il paese. Stojanović concludeva il suo testo in maniera ottimistica, fiducioso nella «vittoria della ragione» e nell'imminente consolidamento del suo alfabeto jugoslavo (733).

In un numero successivo della rivista *Život i rad* (118), apparso nello stesso anno, venne pubblicato un breve testo di commento alla proposta di Stojanović, scritto dal dottor Vojislav Kujundžić,¹² e intitolato «Jugoslovenska latinica» (La latinica jugoslava). Esso si apriva con la seguente affermazione:

Presso una nazione con un'identica lingua nazionale, in uno stato nazionale con un unico sovrano e con un nome comune, come nel caso della nostra nazione jugoslava composta da tre popoli, si è avvertita fin da subito e si sente in maniera crescente giorno dopo giorno la forte necessità di un unico sistema di scrittura. (Kujundžić 1934, 862)¹³

L'autore giustificava questa affermazione innanzitutto su basi pratiche, illustrando l'enorme semplificazione nella stampa che sarebbe risultata dall'uso di un unico sistema di scrittura, con i relativi benefici a livello di alfabetizzazione popolare. La proposta di Kujundžić era simile a quella lanciata da Stojanović qualche mese prima, sebbene egli chiamasse questo alfabeto «latinica jugoslava», una scelta giustificata dall'autore attraverso l'osservazione dell'esistenza al mondo di alfabeti su base latina diversi, ognuno con le proprie peculiarità: essi avrebbero potuto quindi essere meglio distinti dall'uso di un marcatore nazionale (864). Anche l'alfabeto di Kujundžić consisteva in una combinazione di caratteri latini e cirillici, seppure ap-

je mesecima morao da uči, [...] Zašto bismo mi tapkali u mestu, zašto mi ne bismo naše dve azbuke zamенили једном Југословенском, без и једног туђег слова, када је то од првокласног значаја и интереса целине нашега народа, а несравњено лакше и прихватљивије?».

12 Dalle scarse informazioni disponibili, sappiamo che fu medico e fondatore del Rotary Club di Belgrado nel 1927.

13 «Код народа са једним истоветним народним језиком, у једној народној држави, са једним Владаром и са једним заједничким именом, као што је нас троплемени народ у Југославији, осетила се одмах и сваким даном све више осећа потреба за једном писменицом».

parisse piuttosto differente da quello creato da Stojanović. Pure qui, i sette caratteri comuni tra i due alfabeti furono mantenuti, e tutti i digrammi furono rimossi da quello latino: <dž>, <lj> e <nj>. In contrasto con l'alfabeto jugoslavo, tutti i caratteri che presentavano dei segni diacritici, ovvero <ž>, <č>, <ć> e <š>, furono eliminati e sostituiti con quelli corrispondenti presi dal cirillico, insieme ad altre cinque lettere latine. Nel complesso, ecco come si presentava la *latina* jugoslava di Kujundžić:

a	b	v	g	d	ђ	e	ж	z	i	j	k	л	љ	m
н	њ	o	п	p	s	t	ћ	u	f	x	c	џ	ч	ш

4 L'alfabeto jugoslavo di Pavle Ž. Radivojević

Il terzo autore a proporre una soluzione alfabetica mista fu il serbo Pavle Ž. Radivojević, nel suo breve volume intitolato *Ćirilica - Latini-ca? Ili Ćirilica u Latinici?* (Cirillico - latinica? O cirillico nella latinica?), stampato a Belgrado nel 1934. Basando le sue affermazioni su motivazioni simili a quelle degli altri due autori, Radivojević sosteneva come il latino croato e il cirillico serbo provenissero «dalla stessa fonte», avendo essi condiviso simili vicissitudini a livello storico. In riferimento al suo uso nel mondo contemporaneo, Radivojević individuava alcune carenze dell'alfabeto latino e constatava soprattutto l'inadeguatezza dell'alfabeto inglese, definendolo come «illeggibile». L'autore ricordava i tentativi fatti già nel 1818 per la realizzazione di un alfabeto unificato universale, riferendosi a quello sviluppato su base latina dallo storico e orientalista francese Comte de Volney, fondato sul principio della corrispondenza univoca tra fonema e grafema.¹⁴ L'autore esprimeva dunque più concretamente il suo pensiero:

Lo scopo di questo lavoro è di contribuire a far sì che, nei territori abitati da tre popoli della stessa nazione in cui le aspirazioni centinarie si sono recentemente tradotte in viva realtà, dopo quasi due decenni di vita libera, si giunga all'impiego generale quotidiano di un solo alfabeto. È possibile, finalmente, risolvere tale questione attraverso il compromesso in modo tale che nella costruzione del nostro futuro alfabeto partecipino in egual misura sia cirillico che latino. (Radivojević 1934, 18)¹⁵

¹⁴ In merito a ciò, si veda Lepsius 1863, 20-2.

¹⁵ «Смер је овог рада да допринесе да на територији насељеној трима племенима једног народа на којој су недавно вековне аспирације прешле у живу стварности дође, после скоро две деценије слободног живота, до опште свакидашње употребе

A detta del suo inventore, la possibilità di applicare la nuova soluzione di scrittura era realistica, poiché diversi esempi attestavano il fatto che tale contatto fra i due alfabeti già avveniva nella pratica comune, soprattutto nell'uso privato, talvolta a livello involontario (19). A suo avviso, bisognava riconoscere il potenziale valore simbolico, ideologico e unificante del nuovo alfabeto jugoslavo:

È ben noto il fatto che ci sono dei serbi che vanno a vivere in mezzo ai croati e cominciano ben presto a scrivere in alfabeto latino, e così allo stesso modo ci sono croati e sloveni che abitano in mezzo ai serbi e iniziano a servirsi del cirillico. Come spiegare tutto ciò? Con il fatto che per il nostro uomo perspicace, che sia serbo o croato o sloveno, l'alfabeto che fino al 1918 è stato simbolo della nazione e della religione, oggi in Jugoslavia è diventato ciò che un sistema di scrittura effettivamente è: una convenzione sociale. Tale adattamento è in qualche modo un atto di convivialità e di raffinatezza. (19)¹⁶

Tuttavia, aggiungeva Radivojević, sarebbe stato ingenuo presupporre che i croati che utilizzavano il cirillico tra i serbi avrebbero dimenticato le loro origini croate. Era comprensibile che i croati fossero affezionati al loro alfabeto e lo considerassero il loro sistema di scrittura nazionale, e d'altra parte che i serbi amassero il loro cirillico, radicato nel «patriottismo locale, tribale e territoriale» (22). Inoltre, era molto difficile per i serbi accettare l'abbandono di un alfabeto da loro considerato come il «più perfetto di tutti gli altri alfabeti latini e slavi». ¹⁷ Radivojević presentava dunque la sua proposta concreta di un alfabeto misto, selezionando diciotto caratteri dall'alfabeto cirillico e diciassette dall'alfabeto latino. Tredici di questi caratteri erano in cirillico, dodici erano in latino e cinque erano caratteri comuni. Questo alfabeto consisteva nelle seguenti lettere, che apparivano in un ordine piuttosto inusuale:

само једне азбуке. Могућно је, најзад, да се ово питање реши компромисом тако да у изградњу будуће наше заједнице азбуке равноправно учествују латиница и ћирилица».

16 «Без сумње није непознато да има ортодоксних Срба који дођу међу Хрвате и почну убрзо писати латиницом, а тако исто има Хрвата и Словенаца који се настану међу србима и почну се служити ћирилицом. Чим се то објасњава? Тиме да је за увиђаван нашег човека, био он Србин или Хрват или Словенац, писмо које му је до 1918 г. било симбол нације и вере, данас у Југославији постало оно што писмо и јесте: друштвена конвенција. То прилагођавање је донекле акт друштвености и углађености».

17 Troviamo qui un interessante riferimento all'idea di perfezione dell'alfabeto serbo, che sembrava essere piuttosto diffusa all'epoca (cf. Živaljević 1935, 3, 10)

a	b	c	d	e	f	g	x	i	j	k	l	m	n	o
п	р	с	т	у	в	з	љ	њ	ђ	ћ	ж	ч	ш	џ

L'autore concludeva il suo breve volume affermando la necessità di stabilire una nuova tradizione per marcare l'inizio di una storia jugoslava comune, «rispettando le tradizioni che sono da onorare, e sostituendo quelle che sono da rimuovere. Sappiamo che introdurre un alfabeto per tutti significa collocare una grossa pietra nelle fondamenta del nostro nuovo edificio» (22) Tale era l'ottimistica speranza di un convinto sostenitore dell'unità spirituale, politica e culturale dei popoli che componevano l'allora Regno di Jugoslavia.

5 Alcune reazioni alle proposte della nuova scrittura mista: Arandelović e Belić

In un numero della rivista *Život i rad* uscito nel 1934, apparve un commento di D. Arandelović, medico di professione, relativo alle proposte riguardanti la creazione di un sistema di scrittura misto, in un articolo intitolato «Dva mišljenja o jugoslovenskoj azbuci» (Due opinioni sull'alfabeto jugoslavo). Il testo si presentava particolarmente critico e tagliente, al punto che gli stessi redattori della rivista dichiararono in una nota di aver deciso di pubblicarlo pur non essendo d'accordo con i pensieri ivi espressi, tra cui i seguenti:

L'alfabeto jugoslavo! Certo, quando gli 'integralisti jugoslavi' abbandonano il nome serbo, croato e sloveno, quando non vogliono saperne più della lingua serbocroata o slovena, bensì solo di quella 'jugoslava', quando già è in atto la scrittura della storia della Jugoslavia e non quella secolare dei serbi, croati e degli sloveni [...] allora perché non abbandonare il cirillico e il latino per creare un 'alfabeto jugoslavo'? (Arandelović 1934, 861)¹⁸

Il testo di Arandelović fu seguito, un anno dopo, da un commento dell'eminente linguista e filologo Aleksandar Belić, intitolato «Nova azbuka» (Il nuovo alfabeto) e pubblicato sul primo numero della rivista *Naš Jezik*. Belić introduceva il delicato tema della controversia ricordando come il XIX secolo fosse stato segnato dalla lotta per il raggiungimento degli importanti accordi sulla lingua letteraria comune

¹⁸ «Југословенска азбука! Наравно, кад 'интегрални Југословени' напуштају српско, хрватско и словеначко име, кад неће да знају за српско-хрватски и словеначки језик него за 'југословенски', кад се већ пише историја Југославије а не вековна историја Срба, Хрвата и Словенаца [...] онда зашто не бисмо напустили ћилицу и латиницу и створили 'југословенску азбуку'?».

e sull'alfabeto: si era trattato di una «guerra alfabetica», che aveva coinvolto «le due parti della comune nazione» (Belić 1935, 1). Non ci si doveva quindi stupire delle critiche e opposizioni a certe pratiche di scrittura provenienti da membri della società jugoslava. A quel tempo, sebbene un numero non trascurabile di persone conoscesse entrambi gli alfabeti, questo non era certamente un dato di fatto per la maggioranza della popolazione, e di ciò Belić si rammaricava (2).

Tuttavia, la situazione ideale immaginata dal linguista non consisteva nell'istituzione di un nuovo o unico sistema di scrittura, bensì nella trasformazione della situazione di bialfabetismo ufficiale in una di reale bialfabetizzazione, con l'uso equo e spontaneo di entrambi gli alfabeti da parte di tutta la popolazione, attraverso una politica statale che incoraggiasse al massimo l'uso intercambiabile del cirillico e del latino. In tale modo, entrambi i sistemi di scrittura sarebbero stati percepiti da ogni abitante del paese come propri, e ciò avrebbe contribuito a minimizzare i sentimenti patriottici locali.

Belić sembrava tuttavia riconoscere i meriti degli iniziatori della bizzarra soluzione di un alfabeto misto: la «terza via» poteva infatti contribuire a placare le rivalità tra i diversi gruppi etnici, portando a una vittoria «senza vincitori né vinti» (2). Di per sé, sosteneva Belić, una tale proposta non era irrealizzabile, anche perché l'alfabeto latino e quello cirillico condividevano la stessa origine, essendo entrambi discendenti dell'alfabeto greco. Vi era dunque una serie di aspetti positivi indubbi nelle proposte avanzate da Kujundžić, Stojanović e Radivojević. Pur condividendo la necessità di trovare una posizione di compromesso, lo studioso confermava tuttavia come l'alfabeto jugoslavo non potesse incarnare la via più appropriata per la soluzione al problema della presenza contemporanea di due sistemi di scrittura nel paese.

È davvero necessario intraprendere anche nel XX secolo una 'guerra alfabetica' che sopraggiungerebbe inevitabilmente nel momento in cui tentassimo di sostituire il nostro cirillico, che l'intero mondo ha lodato come uno degli alfabeti più perfetti, e il nostro latino, perfettamente [...] corrispondente a esso, con un qualche terzo alfabeto, seppure composto da elementi di entrambi? (3)¹⁹

Secondo lo studioso, parecchie difficoltà sarebbero emerse dall'introduzione di un terzo alfabeto, e queste includevano il non irrilevante allontanamento del popolo da tutte le opere scritte fino ad al-

19 «Да ли треба и у XX веку покретати поново 'азбучни рат', што би неминовно дошло када бисмо нашу ћирилицу, коју је цео свет прославио као једну од најсавршенијих азбука, и нашу латиницу, која би јој [...] у потпуности одговарала - покушали да заменимо нечим трећим, макар и састављеним од елемената обеју?».

lora in latino o in cirillico. Non ci si poteva infatti aspettare che le masse fossero in grado di familiarizzarsi pure con gli altri sistemi di scrittura, una volta imparato l'alfabeto jugoslavo. Tuttavia, come abbiamo visto, i difensori del nuovo alfabeto misto avevano insistito proprio su questo punto, facendo notare come lo sforzo richiesto per imparare i pochi caratteri rimasti di entrambi i sistemi di scrittura risultasse minimo.

Belić concludeva il suo testo rimarcando il valore delle varie proposte per la creazione di un nuovo alfabeto, vedendo in esse uno sforzo lodevole nonché un fattore di sostegno nel cammino di unificazione culturale del popolo jugoslavo. Secondo lo studioso, tentativi simili non si sarebbero arrestati fintantoché l'alfabeto cirillico e quello latino non avessero raggiunto la piena parità nel loro status ufficiale, venendo entrambi «sentiti» dall'intera nazione come genuinamente comuni (3). L'affermazione di Belić rappresentava quindi una piena concretizzazione della necessità di pervenire a una situazione di bialfabetizzazione sia a livello ufficiale che privato, imbarcandosi sulla difficile via della coesistenza ideologica di questi due sistemi di scrittura.

6 La reazione di Danilo A. Živaljević e l'applicazione pratica delle proposte di scrittura

Nel contesto dei dibattiti sulle scelte in ambito di scrittura, Danilo A. Živaljević, scrittore e storico letterario, pubblicò a Belgrado nel 1935 un libretto intitolato *Ćirilica i Latinica* focalizzato sull'analisi delle proposte di alfabeto misto portate avanti da Stojanović, Kujundžić e Radivojević (Živaljević 1935, 28). Questo autore è l'unico a fornire esempi molto concreti sull'applicazione dei tre sistemi di scrittura, riportando ad esempio la trascrizione del nome della capitale jugoslava che in alfabeto latino appare scritto come *Beograd*, e in cirillico come Београд:

- a. Nell'alfabeto di Kujundžić, il nome della città apparirebbe scritto «Beogpad»;
- b. Nell'alfabeto di Stojanović «Беогпад».
- c. Nell'alfabeto di Radivojević «Beogpad».

I famosi versi del poeta Njegoš, 'Бог се драги на Србе разљути/За њихова смртна сагрјешења' in cirillico e 'Bog se dragi na Srbe razljuti/Za njihova smrtna sagrješenja' in latino apparirebbero scritti nei seguenti modi:

- a. Nell'alfabeto di Kujundžić: «Bog se dragi na Spbe razljuti/ Za њихова смртна сагрјешења».
- b. Nell'alfabeto di Stojanović: «Бог се драги на Србе разљути / За њихова смртна сагрјеђења».

- c. Nell'alfabeto di Radivojević: «Bog se drpaga na Spbe pazluti/
Za njihova smptna sagrješenja».

Živaljević sminuiva, a livello sia teorico che pratico, i potenziali benefici di un sistema di scrittura misto in relazione ai problemi che il suo paese doveva affrontare; in particolare, sfatava l'illusione che tale alfabeto sarebbe stato in grado di arrestare «l'antagonismo tribale» presente in molte parti del paese o il suo possibile contributo alla soluzione dei problemi di scrittura e lettura della popolazione anziana. Inoltre, in linea con le dichiarazioni di Belić, Živaljević dichiarava che due alfabeti potevano tranquillamente coesistere, poiché non c'era l'urgenza di risolvere la situazione di digrafia caratteristica del paese:

Due alfabeti non danno fastidio l'uno all'altro [...]. Il cirillico non esclude il latino, così come quest'ultimo il cirillico. Rispettiamo allora sia l'uno che l'altro. Non poniamo nessuno al di sopra dell'altro. Il tempo, che crea e distrugge, risolverà anche la questione dei due alfabeti nel nostro popolo. Indipendentemente dal nuovo alfabeto, che si chiami latinica jugoslava, alfabeto jugoslavo o cirillico nel latino. (34)²⁰

In breve, secondo Živaljević, essendo ormai entrambi gli alfabeti divenuti sistemi di scrittura ufficiali a livello nazionale, era opportuno imparare a rispettarli entrambi, lasciando ognuno libero di decidere quale usare: se il latino o il cirillico.

7 Valutazione critica

Da un punto di vista sociolinguistico, la lingua serbo-croata ha rappresentato un esempio quasi unico al mondo, mettendo in crisi il principio secondo cui una lingua debba venire trascritta mediante un solo sistema di scrittura alla volta (Gelb 1952, 227). Lo stesso vale per la lingua serba contemporanea, nella quale vengono ancora utilizzati entrambi gli alfabeti, sebbene a partire dalla disgregazione della Jugoslavia i documenti ufficiali privilegino in maniera crescente l'alfabeto cirillico. Tuttavia, «se è vero che in generale una lingua sceglie una sola scrittura come mezzo di espressione, non ci sono limitazioni all'uso di una scrittura per un numero qualsiasi di lingue» (228).

²⁰ «Две азбуке једна другој не сметају. [...] Ћирилица не искључује латиницу, као ни ова ћирилицу. Поштујмо и једну и другу. Не истичимо једну над другом. Време, које гради и разграђује, решиће и питање двеју азбука у нашем народу. без обзира на нове азбуке, па ма се оне звале Југословенска латиница, Југословенска азбука или Ћирилица у латиници».

Allo stesso tempo, ipoteticamente non ci sono limitazioni alla possibilità di trascrivere una lingua con un qualsiasi sistema di scrittura, persino uno creato ad hoc e artificialmente a scopi politici, come il cosiddetto ‘alfabeto jugoslavo’ di inizio anni Trenta.

Valutare in maniera critica le proposte di introduzione di un nuovo sistema di scrittura misto nel Regno di Jugoslavia a inizio anni Trenta rappresenta un’impresa ardua, dal momento che tali esempi rappresentano un caso decisamente peculiare, se non addirittura unico a livello linguistico. Nelle teorie sociolinguistiche, soprattutto quelle riferite ai sistemi di scrittura (vedi Sebba 2009), è difficile trovare modelli affini con cui poter trarre dei paragoni o attraverso cui valutare la fattibilità di una simile opera. Di certo, appare opinabile l’idea che un sistema di scrittura creato dalla combinazione in uguale proporzione di caratteri provenienti da due alfabeti rappresentanti due gruppi etnici parlanti la stessa lingua²¹ (o due varianti della stessa lingua) potesse o possa rivelarsi in grado di minimizzare le differenze culturali legate alle parallele influenze della cosiddetta *Slavia orthodoxa* e della *Slavia romana* (si veda Picchio 1991) nel corso dei secoli, contribuendo alla pacifica coesistenza fra serbi, croati e gli altri gruppi etnici presenti nel paese. Così come appaiono decisamente dubbie le capacità di tale sistema di scrittura ibrido in termini di facilitazione all’apprendimento degli alfabeti cirillico e latino stessi nel contesto di lettura, dal momento che, se la riforma proposta fosse stata realizzata, questi due sistemi di scrittura preesistenti si sarebbero trovati ‘marginalizzati’ a livello di dominio d’uso.

Pur riconoscendo le mancanze e le ovvie difficoltà nell’implementazione di tali proposte, è tuttavia importante riconoscere agli autori la buona volontà di aver immaginato una maniera concreta e originale per unificare la lingua a livello alfabetico, risolvendo l’ambiguità della compresenza fra cirillico e latino.

Tali proposte si legavano in maniera decisa agli ideali di unificazione universalista a livello linguistico che si sviluppavano in altre parti d’Europa e del mondo nel periodo interbellico, di cui l’esperanto (che in quegli anni conosceva grande sviluppo e successo), il *Novial* e altre lingue artificiali incarnano un importante paradigma. Allo stesso tempo, altri ‘esperimenti’ unificanti, questa volta a livello di scrittura erano rappresentati dai tentativi di introdurre l’alfabeto latino in Cina (Henze 1977, 393), in Giappone (Wellish 1978, 89-93), così come altrove (si veda l’opera *L’adoption universelle des caractères latins*, 1934), seppure, da quanto ci risulta, nessun altro si spinse così lontano da teorizzare un nuovo alfabeto misto volto a risolvere manifestazioni grafiche di ‘dualismo culturale’.

²¹ Tutti e tre gli autori si riferiscono alla lingua definendola come ‘lingua jugoslava’.

8 Conclusioni

In questo articolo, ho inizialmente presentato il contesto delle ideologie e scelte di scrittura (relative all'alfabeto latino e cirillico) attive nel Regno di Jugoslavia durante il periodo interbellico, concentrandomi sul bialfabetismo e sulla bialfabetizzazione come politiche ufficiali del paese. Ho in seguito analizzato le particolari proposte di riforma alfabetica che presero la forma del cosiddetto 'alfabeto jugoslavo', un sistema di scrittura artificiale e misto, contenente i caratteri di entrambi gli alfabeti. Sebbene mai prese in esame dalla politica ufficiale, queste proposte trovarono eco nei dibattiti degli intellettuali dell'epoca, come dimostrato dagli articoli del filologo Aleksander Belić, di D. Arandelović e di Danilo A. Živaljević qui brevemente illustrati. In tale modo, ho voluto dimostrare come, nelle sue tre varianti, l'alfabeto jugoslavo, sebbene solo ipotizzato e mai implementato, costituisca una preziosa e inusuale testimonianza del grado in cui le idee di fraternità e unità giunsero a permeare le concezioni linguistiche e di scrittura dell'epoca della prima Jugoslavia.

La mancanza di informazioni più dettagliate sulle vite degli studiosi coinvolti nelle proposte di riforma alfabetica ha reso impossibile un'analisi approfondita della genesi e del contesto di sviluppo dell'idea di sistema grafico ibrido, nonché una considerazione degli eventuali contatti fra gli autori stessi. Permangono pertanto ancora dei punti non chiari nella vicenda, e la speranza è che nuove fonti possano essere reperite in futuro, le quali ci permettano di ricostruire con maggiore precisione il contesto di 'idealismo alfabetico' interbellico, e gettino luce su ulteriori dettagli di questa affascinante vicenda rimasta finora completamente sconosciuta al pubblico.

Bibliografia

- Arandelović, D. (1934). «Dva mišljenja o jugoslovenskoj azbuci» Два мишљења о југословенској азбуци (Due opinioni sull'alfabeto jugoslavo). *Život i rad* Живот и рад (Vita e lavoro), VII(118), 861-2.
- Baglioni, D.; Tribulato, O. (2015). «Contatti di lingue - Contatti di scritture: considerazioni introduttive». Baglioni D.; Tribulato O. (a cura di), *Contatti di lingue-contatti di scritture. Multilinguismo e multigrafismo dal Vicino Oriente Antico alla Cina contemporanea*. Venezia: Edizioni Ca' Foscari, 9-38. <http://doi.org/10.14277/6969-061-7/FMM-9-0>.
- Banac, I. (1984). «Main Trends in the Croat Language Question». *Aspects of the Slavic Language Question*. Vol. 1, *Church Slavonic—South Slavic—West Slavic*. Edited by R. Picchio and H. Goldblatt. New Haven: Yale Concilium on International and Area studies, 189-259.
- Belić, A. (1935). «Nova azbuka» Нова азбука (Il nuovo alfabeto). *Naš Jezik* Наш језик (La nostra lingua), III(1), 1-3.

- Bugarski, R. (1997). *Jezik od mira do rata* (La lingua dalla guerra alla pace). Beograd: Slovoğraf.
- Bunčić, D. (2016). «A Heuristic Model for Typology». Bunčić, D.; Lippert, S.L.; Rabus, A. (eds), *Biscriptality: A Sociolinguistic Typology*. Heidelberg: Universitätsverlag Winter, 51-72.
- Collin, R.O. (2011). «Revolutionary Scripts». Morris, M. (ed.), *Culture and Language. Multidisciplinary Case Studies*. Frankfurt am Main: Peter Lang, 29-67.
- Dale, I.R.H. (1980). «Digraphia». *International Journal of the Sociology of Language*, 26, 5-13.
- Feldman, L.; Barac-Cikoja, D. (1996). «Serbo-Croatian: A Biscriptal Language». Daniels, P.; Wright, W. (eds), *The World's Writing Systems*. New York: Oxford University Press, 769-72.
- Gelb, I.J. (1952). *A Study of Writing*. Chicago: University of Chicago Press.
- Greenberg, R. (2004). *Language and Identity in the Balkans*. New York: Oxford University Press.
- Henze, P.B. (1977). «Politics and Alphabets in inner Asia». Fishman, J. (ed.), *Advances in the Creation and Revision of Writing Systems*. The Hague: Mouton, 371-420.
- Jespersen, O. (1934). «Introduction». Société des Nations 1934, 13-26.
- King, R.D. (2001). «The Poisonous Potency of Script: Hindi and Urdu». *International Journal of the Sociology of Language*, 150, 43-59.
- Kujundžić, V. (1934). «Jugoslovenska latinica» Југословенска латиница (La latinica jugoslava). *Život i rad* Живот и рад (Vita e lavoro), VII(118), 862-4.
- Lepsius, K.R. (1863). *Standard Alphabet for Reducing Unwritten Languages*. London: Williams & Norgate.
- Peregud, T.K. Pogoreljski (1924). *Album šrifta: ćirilica i latinica* (Album di caratteri: cirillico e latino). Zemun.
- Petrucchi, A. (2002). *Prima lezione di Paleografia*. Roma-Bari: Laterza.
- Picchio, R. (1991). *Letteratura della Slavia ortodossa (IX-XVIII sec.)*. Bari: Dedalo.
- Pravopisno upustvo (1929). *Pravopisno uputstvo za sve osnovne, srednje i stručne škole u Kraljevini S.H.S* (Manuale ortografico per tutte le scuole primarie, medie e professionali nel Regno dei Serbi, Croati e Sloveni). Ministarstvo Prosvete kraljevine SHS. Beograd: Državna štamparija.
- Radivojević, P.Ž. (1934). *Ćirilica - Latinica? Ili Ćirilica u Latinici* Ћирилица—Латиница? Или Ћирилица у Латиници (Cirillico - latino? O cirillico nel latino). Beograd: Privrednik.
- Rivlina, A. (2016). «Global English-Related Digraphia and Roman-Cyrillic Biscriptal Practices». *Procedia—Social and Behavioral Sciences*, 236, 207-12.
- Sebba, M. (2009). «Sociolinguistic Approaches to Writing Systems Research». *Writing Systems Research*, 1(1), 35-49.
- Selvelli, G. (2018). «L'impatto delle ideologie sovietiche di latinizzazione nei dibattiti bulgari del periodo interbellico: l'inchiesta della rivista Bălgarska Kniga». Maurizio, M.; Tomelleri, V.S. (a cura di), *Rivoluzione visiva attraverso visioni rivoluzionarie: alfabeti, cinema e letteratura in URSS*. Torino: Dipartimento di Lingue e Letterature Straniere e Culture Moderne, Università di Torino, 49-62.
- Skerlić, J. (1914). «Anketa o južnom ili istočnom narečju u srpsko-hrvatskom književnosti» Анкета о јужном или источном наречју у српско-хрватској књижевности (Indagine sul dialetto meridionale o orientale nella letteratura serbo-croata). *Srpski književni glasnik* Српски књижевни гласник (Gazzetta letteraria serb), XXXII(2), 114-25.

- Société des Nations. (1934). *L'adoption universelle des caractères latins*. Institut international de coopération intellectuelle, Paris: Librairie Stock.
- Stefanović, S. (2015). «O javnoj upotrebi jezika i pisma» О јавној употреби језика и писма (Sull'utilizzo ufficiale di lingue e alfabeti). Kusturica, E. (ed.), *Odjeljenje za Istoriju* (Dipartimento di storia). Višegrad: Andrićev Institut, 109-38.
- Stojanović, B. (1934). «Jugoslovenska Azbuka» Југословенска азбука (L'alfabeto jugoslavo). *Život i rad* Жизнь и рад (Vita e lavoro), VII(116), 724-33.
- Stošić, P. (1931). «Nekoliko reči o potrebi jednog pisma» Неколико речи о потреби једног писма (Alcune parole sull'utilizzo di un alfabeto). *Letopis Matice Srpske* Летопис Матице Српске (Annuario della Matica Srpska), CV(327), 154-8.
- Udicki, J. (1931). *Način učenja latinice posle ćirilice i ćirilice posle latinice* (Metodo per imparare il latino dopo il cirillico e il cirillico dopo il latino). Novi Sad: Štamparija učiteljskog d. Natošević.
- Vaillant, A. (1934). «Yugoslavie». Société des Nations 1934, 184.
- Wellish, H. (1978). *The Conversion of Scripts. Its Nature, History, and Utilization*. New York; Chichester; Brisbane; Toronto: John Wiley & Sons.
- Zima, P. (1974). «Digraphia: The Case of Hausa». *Linguistics*, 124, 57-69.
- Živaljević, D.A. (1935). *Ćirilica i Latinica* (Cirillico e latino). Beograd: Privrednik.

К этимологии праславянского обозначения ‘раны’

Artem Trofimov

RANEPA (The Russian Presidential Academy of National Economy and Public Administration)

Abstract The paper is devoted to a critical analysis of the traditional etymology of the Proto-Slavic word with the meaning ‘wound’ (OCS ‘рана’ πληγή, μάστιξ; Bulgarian ‘ра̀на’; Serbocroatian ‘ра̀на’, Russian ‘рана’ etc.). Researchers usually compare it to OInd. *vraṇá-* ‘wound, ulcer’ and Alb. *varrë* ‘wound’. It is demonstrated that this etymology is unacceptable from a formal point of view. Therefore, it is proposed to return to the idea of P. Persson who connected the Proto-Slavic ‘wound’ to the Proto-Indo-European root **h₂erH-* ‘to destroy’ and the following cognates in other branches of PIE: Hett. *ḡarrai* ‘grinds’; Toch. AB *ār*, pret. B *āra* ‘to stop, to cease’ and Proto-Slavic **oriti* ‘to destroy, to ruin’. In this case the Proto-Slavic reconstruction is **ōrna*.

Keywords Proto-Slavic. Etymology. Schwebeablaut. Cognate. Old Indian. Albanian.

Summary 1 Введение. – 2 Связь с древнеиндийской и албанской формами. – 3 Предлагаемые внутриславянские соответствия. – 4 Альтернативная этимология.



Peer review

Submitted	2021-12-11
Accepted	2022-01-11
Published	2022-02-15

Open access

© 2021 | Creative Commons Attribution 4.0 International Public License



Citation Trofimov, A. (2021). “On Etymology of Proto-Slavic Word for ‘Wound’”. *Balcania et Slavia*, 1(2), 225-238.

1 Введение

Существует некоторое количество славянских этимологий, которые являются общепризнанными, зарегистрированы во всех авторитетных этимологических словарях славянских языков, но при этом существуют значительные проблемы с предлагаемыми сопоставлениями, как формальные, так и семантические. В статье пойдет речь об одном из таких случаев.

Праславянское слово, обозначающее 'рану', надежно реконструируется на основании следующих языковых фактов:

рус., укр., блр. *rána*, др.-русск., ст.-слав. *рана* πληγή, μάστιξ (Супр.); болг. *rána*; сербохорв. *rāna* 'рана; зло, невзгода; высеченное место на дереве, из которого вытекает сок', словенск. *rána*, чеш. *rána* 'рана; удар; несчастье', слвц., польск., в.-луж., н.-луж. *rana* 'рана'. (ЭСРЯ III, 441-2)

Наиболее распространенная традиционная этимология предполагает сближение с др.-инд. *vraṇá-* m., n. 'рана, язва' и алб. *varrë*, pl. *varrë* 'рана' (IEW, 1163-4). Она была предложена Э. Лиденом в работе, специально посвященной развитию начальных кластеров **ur-* и **ul-* в балтийских и славянских языках (Liden 1899, 19-20), и поддерживается в большинстве этимологических словарей славянских языков (ЭСРЯ III, 441-2; EJSJ, 747-8; ЕСУМ 5, 23; ЭСБМ 11, 94-5; Brückner 1985, 453; Borys 2005, 510; Machek 1957, 508; Rejzek 2015, 580; Skok 1973, 107; БЕР VI, 180-1; Bezljaj III 1995, 150-1; Králik 2015, 490).

Дальнейшие продолжения на славянской почве, также связываемые с обозначением 'раны', по мнению этимологов, включают такие формы, как рус. *воронка*, диал. *воронá*, также в знач. рыба *Gobius barbatus* (разновидность бычков), укр. *ворóнка*, болг. *vrána* 'пробка, втулка', сербохорв. *vrân* 'запор, втулка', словен. *vrânj* 'то же', чеш. *vřana* 'отверстие', польск. *wrona* 'то же'; далее латышск. *vārna* 'колода с желобками для вития веревок' (ЭСРЯ I, 295). Реконструируемая на их основании праформа **vorna*, параллельная, по мнению этимологов, к **rana* < **urōnā*, относится к праславянскому глаголу **verti* 'вдевать, продевать' и лит. *ver-iù*, *vér-ti* 'пронзать, прокалывать; вдевать, продевать; нанизывать; закрывать, открывать' (IEW 1150-1). Еще одно славянское существительное, иногда приводимое в качестве однокоренного 'ране' – имя **verdъ* 'нарыв, гнойник; вред, повреждение' (IEW, 1163; Rejzek 2001, 527).

На индоевропейском уровне слово увязывается некоторыми исследователями с и.-е. **uer-* 'разрывать, поцарапать', например, такая трактовка предлагается в этимологическом словаре польского языка В. Боряща, этимологическом словаре словенского

языка Л. Кралика и этимологическом словаре чешского языка И. Рейзека (Borys 2005, 510; Králík 2015, 490; Rejzek 2001, 527). Можно видеть, что эта идея является до сих пор актуальной – все процитированные источники вышли в последние два десятилетия.

Следует отметить, что некоторые из авторов этимологических словарей славянских языков называют обсуждаемую этимологию праславянского слова 'рана' сомнительной, среди них А. Брюкнер и авторы этимологического словаря украинского языка (Brückner 1985, 453; ЕСУМ 5, 23). Небезынтересен тот факт, что данное слово отсутствует в этимологическом словаре праславянского языка Р. Дерксена (Derksen 2008), что с большой долей вероятности свидетельствует о том, что автор сомневался в обоснованности внешних сравнений, приводимых в литературе для обсуждаемого слова, и, соответственно, в реконструкции **rana* < **rōnā*, обычно принимаемой в других этимологических словарях славянских языков. В то же время Р. Дерксен в своей работе отмечает другое надежно восстанавливаемое праславянское образование со значением 'рана', ср. **ězva* f. (a) 'wound' (Derksen 2008, 155); можно еще обратить внимание на синонимичное праслав. **strǫpъ* m. (c) 'scab', такие продолжения которого, как др.-рус. *строупъ*, имеют ряд значений: 'рана, особенно гнойная; корка на заживающей ране; труп; грех' (Derksen 2008, 470; СДРЯ XI-XVII (28), 199).

Несмотря на в целом широкое и практически безоговорочное принятие данной этимологии в славистической этимологической литературе, ее нельзя назвать безупречной.

2 Связь с древнеиндийской и албанской формами

Во-первых, сравнение славянских данных с древнеиндийскими и албанскими имеет очевидную формальную слабость. Для древнеиндийского *vraṇá-* m.n. 'рана; язва' может быть реконструирована индоевропейская праформа **urēno-*, поскольку для праформы **urōno-* должен был бы действовать закон Бругманна, в результате чего закономерным результатом в древнеиндийском должна была бы быть незасвидетельствованная форма # *vṛāṇá-*. Аналогичный результат, естественно, должна была дать праформа **rōno-* с долгим корневым вокализмом. Далее, необходимо предполагать наличие корня **ur-* в нулевой ступени и суффикса **-ēno-* (EWAia III, 482). Подобное обстоятельство, безусловно, ослабляет сопоставление с праславянским словом, обозначающим рану. Еще один аргумент в пользу отведения рассматриваемой этимологии – тот факт, что слово в санскрите было засвидетельствовано достаточно поздно; в связи с этим М. Майрхофер считает слово не объясненным убедительно (EWAia III, 482); ранее

он же относился к обсуждаемой этимологии более позитивно, хотя и отмечал, что сопоставление со славянскими и албанскими данными носит предположительный характер (KEWA III, 277).

Что касается алб. *varrë*, pl. *varrë* 'рана', то данное слово обсуждается как в этимологическом словаре албанского языка Б. Демирая, так и в этимологическом словаре В. Орла (Demiraj 1997, 410; Orel 1998, 495). Б. Демирай отмечает, что семантически обсуждаемая этимология (сравнение со славянскими и древнеиндийскими формами) привлекательна, но указывает на сложность обоснования праформы, в нотации автора **(s)uor(H)-en-* (Demiraj 1997, 410). При такой трактовке праформа **uor(H)n-eh₂* может быть понята как деривация, переход в **ā*-основу женского рода. Непосредственные причины и механизм для албанского материала не ясны. В. Орел вообще считает имя *varrë* производным от глагола *vras*, аор. *vrava* 'убивать', который, в свою очередь, реконструируется им как **awa-rautja-* с вторичным глагольным суффиксом *-as* и сравнивается далее с лат. *ruō, -ere* 'рушиться, низвергаться, падать'; также в албанском языке имеется слово *vrajë*, pl. *vraja*, напрямую производное от *vras* (Orel 1998, 514-15). Конечно, если не самое вероятное предположение В. Орла справедливо, сравнение с албанским также неправомерно; более того, албанский начальный кластер *vr-* в данном случае не является отражением и.-е. **ur-*. Но даже в случае принятия сопоставления ст.-слав. рана ~ др.-инд. *vraṇá-* м.п. 'рана; язва' ~ алб. *varrë* 'рана' нельзя не согласиться с Б. Демираем: праалбанская праформа восстанавливается с сомнением. Реконструкция **uornā*, однозначная для праалбанского, представляет собой третий вариант основы по сравнению с предполагаемыми для праславянского **urōnā* и для индоарийского **ureno-*. Напрямую этот третий вариант соответствует праслав. **vorna* 'отверстие', но семантически такое сравнение выглядит проблематично, если учесть, что внутри славянского существительное **vorna* связано с глаголом **verti* 'вдевать, продевать', а албанское слово *varrë* внутри албанского – с глаголом *vras* 'убивать'.

3 Предлагаемые внутриславянские соответствия

Наконец, стоит уделить особое внимание внутриславянским соответствиям, которые предлагаются для ст.-слав. 'рана'.

Наиболее часто приводимое сравнение с **vorna* 'отверстие' представляется семантически возможным. Например, можно привести пример полисемии 'отверстие, дыра / рана' из тунгусо-маньчжурских языков: так, эвенкийское слово *haṇar* характеризуется полисемией 1) 'дыра, отверстие'; 2) 'щель'; 3) 'дупло'; 4) 'яма'; 5) 'пещера'; 6) 'рана' (Дубровская 2014, 215; ССТМЯ 1975,

62). Более того, семантика сербохорватского слова *rāna*, имеющего одно из значений 'высеченное место на дереве, из которого вытекает сок' (РСКЖ V, 401), может указывать на связь исходной семантики с 'отверстием', хотя не менее вероятно позднее вторичное метафорическое развитие. Однако формальное сведение праформы **vorna*, праславянскую акцентную парадигму которой по имеющимся рефлексам установить несколько проблематично, и **rāna* < **urōnā*, с акцентной парадигмой *a*, представляется сомнительным. Последняя в таком виде может объясняться на праиндоевропейском уровне только из **uroH-neh₂*. Вид корня **uroH-* возможен только как вторичный вариант по отношению к корню **uerH-*, полученный в результате швебеаблаута. Литовский глагол *veriù, vérti* 'пронзать, прокалывать; вдевать, продевать; нанизывать; закрывать, открывать', соответствующий праслав. **ver-ti* 'вдевать, продевать', ср. ст.-слав. *въврѣти* 'всунуть' и другие родственные формы (ЭСРЯ I, 293), действительно указывает на ларингал (Derksen 2015, 498). У Ю. Покорного это корень, записанный в устаревшей нотации без schwa primum **uer-*, который глоссируется исследователем как 'binden, anreihen, aufhängen' (IEW, 1150-1). В таком случае следует предполагать, что в праязыке существовали два варианта обсуждаемого корня, **uerH-* и **uroH-*, производные которых имели, с одной стороны, значения 'вдевать, продевать' и далее 'отверстие', а с другой – 'рана', возможно – 'повреждать', причем второй вариант представлен только в славянских языках. Подобный сценарий представляется маловероятным. Более того, механизм швебеаблаута в праязыке, как показано в новейшем исследовании, связан вовсе не с такими процессами, как образование вторичных **CReC*-вариантов корня от формы с нулевой ступенью **CṛC-*, а с разного рода деривационными процессами (Ozoliņš 2015, 147). Образования с суффиксом **-neh₂* очевидным образом не относятся к словам, которые обычно образуются от тех или иных архаичных праязыковых имен с нестандартным аблаутом; несколько позже будет рассмотрено некоторое количество праславянских образований, унаследованных от индоевропейского праязыка и обычно образованных от глагольных корней. Наиболее близкий пример, касающийся **-no*-образований – прилагательное (причастие) **dμη₂-nó-* 'построенный' (к состоянию корня II **demh₂-* 'строить'), отраженное с швебеаблаутом в греч. дор. *νεο-δμῆτος* 'недавно построенный'; от него произведена субстантивированная основа **dme₂-no-*, засвидетельствованная в наиболее архаичном слое авестийского языка в виде *dātāna-* 'строение, здание' (Steer 2015, 49; Ozoliņš 2015, 142).

Сравнение с существительным **verdъ* 'нарыв, гнойник; повреждение, вред', реконструируемого на основании таких форм, как др.-русс. *вередъ* 'рана, нарыв', ст.-слав. *врѣдъ* (др.-греч. βλάβη, μῶλωψ), русск. *вред* (из церк.-слав.), укр. *véred*, болг. *вредá*

‘вред’, сербохорв. *vrùjed*, словенск. *vrèd* (род. п. *vréda*) ‘повреждение’, чешск. *vřed*, словацк. *vred*, польск. *wrząd* ‘гнойник’, в.-луж. *bród*, н.-луж. *ród* (ЭСРЯ I, 295), является относительно слабым: в этой праформе наблюдается корневое расширение, что снижает вероятность сближения. Для **verdъ* гораздо более вероятно традиционное сравнение с латышск. *ap-virde* ‘нарыв’ и др.-в.-нем. *warza*, др.-исл. *warta*, нов.-в.-нем. *Warze* ‘бородавка’ (ЭСРЯ I, 295; РЭС 6, 273-4). Кроме того, акцентуация укр. *véred* и отдельных русских диалектных форм *véred* со значением ‘нарыв, прыщ’ указывает на отсутствие ларингала в праформе. Акцентная парадигма *a* в словенском языке и ряде русских диалектов вторична, соответственно, исходна акцентная парадигма *b* (Булатова, Дыбо, Николаев 1988, 55-6). Более подробный набор языковых данных, на основании которого реконструируется праслав. **verdъ*, с вниманием в первую очередь к восточнославянским диалектным формам, представлен в (РЭС 6, 273-4).

Последнее распространенное утверждение, которое отмечается в славистической литературе – наличие праиндоевропейского корня **uer-* со значением ‘разрывать, царапать’ (Boryś 2005, 510; Králik 2015, 490; Rejzek 2001, 527). В отношении данной гипотезы стоит отметить, что глагольный корень в таком виде и с такой семантикой надежно не засвидетельствован в индоевропейских языках. В “Этимологическом словаре индоевропейского глагола” есть пять корней вида **uer(H)-*, и не один из них не подходит по семантике: **uer⁻¹* ‘aufhalten, (ab)wehren’; **uer⁻²* ‘laufen’; **uer⁻³* ‘beobachten, wahrnehmen’, а также **uerH-* ‘heiß sein’ и **uerh_i-* ‘sagen’ (LIV², 684-6; 689-90). Представляется возможным только предполагать его наличие, основываясь на сомнительном сближении именных форм в славянских, индоиранских и албанском языках. Соответственно, подобное предположение никак не проясняет картину и не может служить аргументом в пользу традиционной этимологии праславянского обозначения ‘раны’.

Все приведенные соображения заставляют отказаться от сближения ст.-слав. *рана*, рус. *рана* с др.-инд. *vṛaṇá-* м., п. ‘рана, язва’ и алб. *varrë*, пл. *varrë* ‘рана’, отринуть сравнение с такими внутриславянскими параллелями, как **vorna* ‘отверстие’ и **verdъ* ‘нарыв, гнойник; повреждение, вред’, и искать другие возможности.

4 Альтернативная этимология

На наш взгляд, более удачное объяснение рассматриваемого слова в общих чертах предложил П. Перссон в первом томе своего важного исследования "Работы по индоевропейскому словообразованию" (Persson 1912). Этимология шведского исследователя предполагает сравнение с др.-инд. *īrma-* 'рана' (не засвидетельствовано в ведийский период); лит. *inrù / yrù, irti* 'распадаться; разваливаться; распарываться; расстраиваться (о здоровье)', для которых П. Перссоном реконструируется в духе господствовавших в то время младограмматических представлений "двусложная тяжелая база" **er-* / **erē-*; также автор привлекал для сравнения латышск. *rēta, rēte* 'шрам', лит. *retī-s, rēti-s* 'рана, шрам', ст.-слав. *рѣдѣкъ* 'редкий' и лат. *rārus* 'редкий' (Persson 1912, 278). В современных терминах такие корни восстанавливаются с ларингальным после сонорного. В "Словаре индоевропейского глагола" для обсуждаемого корня предлагается реконструкция **h₂erH-* 'разрушать(ся)' на основании следующих данных: хетт. *ḫarrai* 'размалывает, размельчает'; тох. АВ *ār*, претерит В *āra* 'прекращать, останавливаться'; лит. *inrù / yrù, irti*; рус. *разорю, разорить* (LIV₂ 2001, 271-2).

Для корневого элемента праславянского обозначения раны ученый предполагал, соответственно, ступень **rō-* в корневом элементе. В этом смысле его этимология принципиально сходна с версией Э. Лидена, поскольку праформа **rōnā* отличается от **uḡrōnā* исключительно отсутствием начального согласного, выпадавшего в таких кластерах, как и было показано в (Liden 1899). Этим не исчерпываются недостатки реконструкции П. Перссона: если исходить из современных представлений, аблаут **rē:* **rō* не является стандартным для корней вида **HerH-*¹, в связи с чем также сопоставление с латышск. *rēta, rēte* 'шрам' и ст.-слав. *рѣдѣкъ* 'редкий', лат. *rārus* 'редкий' выглядит непривлекательным. Но есть другая возможность связать слово 'рана' в праславянском с обсуждаемым индоевропейским корнем **h₂erH-*. Хорошо известно, что начальное сочетание *ra-* во всех славянских языках можно также получить из праслав. **ōr-* с акутовой интонацией, ср. праслав. **ōrmē, *ōrmen-, *ōrmo* 'плечо', др.-рус. *рамя*, *рамо* 'плечо', родственное др.-инд. *īrmá-* м. 'рука, передняя часть туши животного', др.-прус. *irmo* 'рука' < и.-е. **HorHmo-* / **H₁Hmo-* (ЭСРЯ III, 440-1; IEW, 332-3). Если оперировать корнем **h₂erH-*, можно реконструировать праслав. **ōrna-* с хорошо засвидетельствованной акутовой интонацией и соответствующей ей акцентной парадигмой *a* < и.-е. **h₂rH-neh₂-*.

¹ Необходимо опять же предполагать швебеаблаут **HerH-*: **HreH-* и одновременно отсутствие воздействия ларингальной окраски на **o*-ступень корня.

Среди дальнейших возможных именных производных от корня $*h_2erH-$ в различных индоевропейских языках можно назвать др.-инд. *áruṣ-* п. 'рана' с другим корневым расширением; традиционно это слово сравнивается с др.-исл. *ørr, err* 'шрам, рубец' (IEW, 338; EWAia I, 113).

Интересно, что и $*ōrna-$ 'рана', и $*ōrmē, *ōrmen-, *ōrmo$ 'плечо' восходят к корням, которые реконструируются в виде $*h_2er(H)-$. На наш взгляд, это разные корни: для названия плеча подходит индоевропейский корень $*h_2er-$ 'скреплять, прилаживать', представленный в том числе в др.-греч. ἀρᾰρίσκω, др.-инд. *ṛtá-* и пр. (LIV₂ 2001, 269-70; IEW, 55-8). Именно такая трактовка для $*ōrmo$ предлагается вслед за В.А. Топоровым в "Этимологическом словаре славянских языков": "сближение приведенных выше образований с греч. ἀρμός 'связь, скрепа, паз', ἄρμα 'конская упряжь', ἀρμόζω, ἀρμόττω 'скреплять, соединять, прилаживать' позволяет предположить, что определяющей для семантики слов была идея соединения, сочленения, связи, откуда значения 'плечо, верхняя часть руки, лопатка'" (ЭССЯ-32, 187). При этом идея возможной связи $*h_2er-$ 'скреплять, прилаживать' и $*h_2erH-$ 'разрушать(ся)' на основании предположения глубинной праязыковой энантиосемии (ЭССЯ-32, 164) представляется менее вероятной.

Соответственно, внутри праславянского слово $*ōrna-$ предпочтительнее связывать не с рус. *воронка* и сходными формами, а с глаголом $*oriti$, который широко представлен в различных языках, в основном с приставками: рус. *разорить, разор*, укр. *розорити, розор*, др.-русс. *оритель* 'разрушитель', ст.-слав. *орити, орѣ* λύειν, καθαιρεῖν, *разорити, сънорити* ἀπορρίπτειν (Супр.), болг. *оря* 'разоряю', сербохорв. *оборити, обориъм* 'сокрушать, ниспровергать, свергать', *разорити, разориъм* 'разорить', чеш. *obořiti* 'обрушиться, напасть' (ЭСРЯ III, 152-3; ЭССЯ-32, 162-5), а также существительным $*ōrtь$ 'война, сражение, рать', очевидно, относящемуся к этому же корню, на что справедливо указано в (Младенов 1941, 557). Судя по тому, что уже в древних славянских языках, как и в современном русском, для обсуждаемого глагола одним из основных является "значение уничтожения денежно-материальных ресурсов противника", ср. в том числе рус. диал. сибир. *иззорить* 'разорить, обокрасть' (Муратов 2015, 135), наиболее вероятно, что праформы $*ōrna$ 'рана, повреждение' и $*oriti$ 'наносить ущерб, разрушать' не осознавались как родственные уже в праславянский период. В дальнейшем с утратой общности формы значения во всех славянских языках значительно разошлись.

Несмотря на то, что этимология П. Персона отвергается в абсолютном большинстве этимологических словарей славянских языков, она в улучшенном виде представляется гораздо более удачной, поскольку формально $*ōrna$ надежно выводится из и.-е. $*h_2e/orH-neh_2$.

Образования с суффиксом **-neh₂* женского рода от корней различного вида широко распространены в праславянском и других группах индоевропейских языков, но есть и примеры сосуществования суффикса **-nā* в праславянском и **-to* и **-ten*-основ в других индоевропейских языках:

Праслав. **luna* 'луна, месяц' (рус. луна), лат. *luna*, -ae f. 'луна' < и.-е. **louksneh₂*- наряду с др.-инд. *ruk-má*- м.п. 'золотое украшение', и прагерм. **leuh-man*- м. 'луч света', ср. др.-исл. *ljómi* м. 'луч света, свечение', фарерск. *ljómi* м. 'свет, блеск', др.-англ. *leoma* м. 'луч света, блеск' и т.д. (ЭССЯ-16, 173-4; IEW, 687-90; Kroonen 2013, 333; EWAia II, 463-4). Все перечисленные формы являются прозрачными производными известного праиндоевропейского корня **leuk-* 'быть светлым; светить', очень хорошо представленного в наиболее древних языках. Строго говоря, в данном случае необходимо вести речь о праязыковом суффиксе **-sneh₂*-. Особенно примечательно, что образования **louk-sneh₂*-, **luk-to*- и **leuk-ten*-, судя по сосуществованию в латыни слов *luna* 'луна' и *lūmen*, -inis п. 'свет', одновременно были представлены в праязыке.

Праслав. **pelena*, **pelna* 'пелена' (рус. пелена, пленка; сербохорв. *пелѐна* 'пелена, оболочка') < и.-е. **pelH-neh₂* наряду с др.-греч. πέλμα п. 'подошва ноги или обуви'; πόλμη f. 'легкий щит'; прагерм. **felma(n)* 'пленка', ср. др.-англ. *æger-felma* м. 'яичная скорлупа', совр. англ. *film*, ср.-нид. *velm*, *vilm* 'пленка, тонкий слой' (IEW, 803-4; Beekes 2010, 1168; Kroonen 2013, 135). В данном случае особенно удивительна славянская праформа **pelena* с мнимым полногласием восточнославянского типа.

Праслав. **pěna* (рус. пена), др.-инд. *phéna*- м. 'пена' < **(s)pHoj-neh₂* наряду с лат. *spūma* f. 'пена, морская пена' и прагерм. **faima(n)*- м. 'пена', ср. англ. *foam* 'пена'. (IEW, 803-4; Beekes 2010, 1168; Kroonen 2013, 135)

Праслав. **vьlna* 'волна' (рус. волна), лит. *vilni-s* 'волна', диал. *vilnià* 'то же', латышск. *vīlnis*, *vīlnis* 'волна', *vilna* 'то же' наряду с др.-инд. *ūrmí*- м., f. 'волна, вал', авест. *varəmi*- 'волна' и прагерм. **walmi*, ср. др.-англ. *wilm* (*wielm*, *welm*, *wālm*, *wylm*), -es м. 'поток, фонтан, кипящая или пенящаяся вода; жар; порывистое движение'. (IEW, 1140-3; Derksen 2008, 537; Derksen 2015, 504; EWAia I, 245)

Любопытно при этом, что балтийских примеров наличия суффикса **nā* < **neh₂* наряду с губными в других языках практически нет, можно привести только достаточно периферийное и достаточно сомнительное в этом отношении образование:

Лит. *karnà* с параллельным вариантом *kardà* 'липовая, ивовая или вербовая кора' < **(s)kor-neh₂* наряду с др.-инд. *cárman-* п. 'шкура, кожа', авест. *čarəman-* 'то же', праслав. **kora* 'кора' и т. д. (IEW, 940; Smoczyński 2021, 632). Данная этимология не может считаться бесспорной, особенно если учесть, что литовское слово достаточно редкое; Э. Френкель полагает, что лит. *karnà* скорее можно сравнивать с литовским же *kirna* 'коряга; бурелом', др.-прусс. *kirno* 'кустарник' и праслав. **koreň* 'корень' (LEW I, 221), что выглядит формально более правдоподобно, но несколько сомнительно с точки зрения семантики.

Чередование суффиксов, содержащих **n* и **m*, может быть прослежено и в других группах индоевропейских языков. Здесь приведем только некоторые случаи; очевидно, тема взаимодействия имены суффиксов с носовыми согласными, а также проблема происхождения **n* из консонантного кластера **mn*, представляющего собой нулевую ступень суффикса **-men-*, заслуживает отдельного рассмотрения на более обширном сравнительном материале:

Др.-инд. *budhná-* м. 'дно, основание; почва' – др.-греч. πυθμήν, -ένοc м. 'дно сосуда, моря и т.д., основание; корень, пень, ствол'. (IEW, 174; EWAia II, 228-9; Beekes 2010, 1255)

Др.-инд. *dānu-* ф. 'ведийский эпитет дождя', этимологически, вероятно, 'река, поток', авест. *dānu-* 'река', осет. *don* 'река, вода' < и.-е. **d^heH-nu-* наряду с др.-греч. δᾱμός, δημός 'жир, тук' (IEW, 175; EWAia I, 719-20; Beekes 2010, 325-6). Данный случай сомнителен, поскольку этимология древнегреческого слова не может считаться ясной. Предлагалось и сближение с алб. *dh-jamë* 'жир, сало', в индоевропейском происхождении последнего сомневается Б. Демирай, которому кажется подозрительным как вокализм, так и начальный *dhj-* < и.-е. **d-*; в крайнем случае албанское слово может быть греческим заимствованием (Demiraj 1997, 161).

Др.-греч. βρέχμός м.; βρέχμα п., βρέγμα п. 'верхняя часть черепа, темя' < **b^hreg^hmo-* (или **mreg^hmo-*) наряду с прагерм. **brag-na-* 'мозг', ср. др.-англ. *brægen*, др.-фриз. *brein* 'мозг'. (IEW, 750; Beekes 2010, 238; Kroonen 2013, 73)

Можно видеть, что наличие образований с **-neh₂* наряду с суффиксами, содержащими **m* в других языках, может наблюдаться в праславянском, как и в других индоевропейских языках (например, итальянских и индоиранских). Данный факт дополнительно доказывает, что предложенная П. Перссоном и уточненная нами этимология допустима.

Еще один аспект, который стоит отметить в связи с сопоставлением праслав. **ǫrna-* 'рана' ~ др.-инд. *īrmá-* п. 'рана' – наличие **o*-ступени в праславянском наряду с нулевой ступенью в древнеиндийском. Как уже было показано, образования с суффиксом **neh₂* с **o*-ступенью могут сосуществовать с именами, имеющими суффикс с губным носовым согласным и нулевую ступень, ср. **louk-sneh₂*- и **luk-mo-* от корня **leuk-* 'быть светлым; светить'. Более того, для некоторых немногочисленных праславянских слов с акутовой интонацией (праиндоевропейским ларингальным), содержащих сонант, характерно присутствие **o*-ступени при балтийских соответствиях с нулевой ступенью, можно привести в качестве примера уже упоминавшиеся праслав. **ǫrmē*, **ǫrmen-*, **ǫrmo* 'плечо' при др.-прус. *irmo* 'рука' и праслав. **koreň* при др.-прус. *kirno* 'кустарник', лит. *kirna* 'коряга; бурелом'.

Проведенное исследование дает возможность заново посмотреть на этимологию П. Перссона для праславянского обозначения 'раны': данное слово следует реконструировать в виде **ǫrna* (а) и связывать с праслав. **oriti* 'разрушать' и **ǫrtь* 'война, сражение, рать'; лит. *inrù / yrù, irti* 'распадаться; разваливаться; распадаться; расстраиваться (о здоровье)'; хетт. *ḫarrai* 'размалывает, размельчает'; тох. АВ *ār*, претерит В *āra* 'прекращать, останавливаться' и с редким др.-инд. *īrma-* 'рана'. Все перечисленные родственные слова восходят к и.-е. **h₂erH-* 'разрушать(ся)'. Отрицание данной этимологии исследователями было связано во многом с некорректной реконструкцией праславянской основы в виде **rana* < **rōpā* и обсуждением преимущественно таких дальнейших внутриславянских связей, как **rědъkъ* 'редкий'. В действительности при принятии предлагаемых поправок она безупречна формально и вполне допустима с семантической точки зрения.

Литература

- БЕР = Български етимологичен речник. Съставили Вл. Георгиев, Ив. Гълъбов, Й. Займов, Ст. Илчев. Т. 1-7. София: Издателство на БАН, 1971-2010.
- Булатова, Р.В.; Дыбо, В.А.; Николаев, С.Л. (1988). Проблемы акцентологических диалектизмов в праславянском // Славянское языкознание. X Международного съезда славистов (София, сентябрь 1988 г. Доклады советской делегации). Москва: Наука.
- Дубровская, Н.В. (2014). "Деривационная структура полисемии: типы производных значений скупского существительного". Сибирский филологический журнал, 1: 211-8.
- ЕСУМ = Етимологічний словник української мови. Гол. ред. О.С. Мельничук. Т. 1-5. Київ: Наукова думка, 1982-2006.
- Младенов, С. (1941). Етимологически и правописен речник на българския книжовен език. София: Книгоиздателство Христо Г. Данов.

- Муратов, Ю.М. (2015). *Генетическая и мотивационная характеристики лексико-семантического поля 'присвоение чужого' в русском языке. Диссертация на соискание ученой степени кандидата филологических наук. Москва.*
- РСКJ = *Речник српскохрватског књижевног језика*. Књ. 1-6. Нови Сад-Загреб: Институт за српскохрватски језик, 1967-76.
- РЭС = Аникин, А.Е. (2007-). *Русский этимологический словарь*. Вып. 1-15. Москва: Ин-т рус. яз. им. В.В. Виноградова РАН; Ин-т филологии Сибирского отделения РАН.
- СДРЯ XI-XVII (1975-). *Словарь древнерусского языка XI-XVII вв.* Вып. 1-30. Москва: Наука – Азбуковник.
- ССТМЯ = Горцевская, В.А.; Колесникова, В.Д.; Константинова, О.А.; Новикова, К.А.; Петрова, Т.И.; Цинциус, В.И.; Бугаева, Т.Г. (1975). *Сравнительный словарь тунгусо-маньчжурских языков*. Т. I-II. Ленинград: Наука. Ленинградское отделение.
- ЭСРЯ = Фасмер, М. (1964-73). *Этимологический словарь русского языка. Перевод с немецкого языка и дополнения О.Н. Трубачева*. Т. I-IV. Москва: Прогресс.
- ЭССЯ = *Этимологический словарь славянских языков. Праславянский лексический фонд*. Под ред. О.Н. Трубачева, А.Ф. Журавлева и Ж.Ж. Варбот. Вып. 1-42. Москва: Наука, 1974-2011.
- Beekes, R.S.P. (2013). *Etymological Dictionary of Greek*. Leiden; Boston: Brill.
- Bezlaj F. (1977-2005). *Etimološki slovar slovenskega jezika*. Knj. I-IV. Ljubljana: Založba ZRC.
- Boryś, W. (2005). *Słownik etymologiczny języka polskiego*. Kraków: Wydawnictwo Literackie.
- Brückner, A. (1957). *Słownik etymologiczny języka polskiego*. Warszawa: Wiedza Powszechna (Przedruk z pierwszego wydania. Kraków, 1927).
- Demiraj, B. (1997). *Albanische Etymologien. Untersuchungen zum Albanischen Erbwortschatz*. Amsterdam; Atlanta (GA): Rodopi.
- Derksen, R. (2008). *Etymological Dictionary of Slavic Inherited Lexicon*. Leiden; Boston: Brill.
- Derksen, R. (2015). *Etymological Dictionary of the Baltic Inherited Lexicon*. Leiden; Boston: Brill.
- EJSJ = *Etymologický slovník jazyka staroslověnského*. Red. E. Havlová, A. Erhart, I. Janyšková. Seš. 1-19. Praha-Brno: Academia-Tribun, 1989-2018.
- EWAia = Mayrhofer, M. (1992-2001). *Etymologisches Wörterbuch des Altindoarischen*. I-III Bde. Heidelberg: Universitätsverlag Winter.
- IEW = Pokorny, J. (1959). *Indogermanisches etymologisches Wörterbuch*. Bern; München: Francke.
- KEWA = Mayrhofer M. (1956-76). *Kurzgefasstes etymologisches Wörterbuch des Altindischen*. Bde. 1-4. Heidelberg: Universitätsverlag Winter.
- Králik, L' (2015). *Stručný etymologický slovník slovenčiny*. Bratislava: VEDA.
- Kroonen, G. (2013). *Etymological Dictionary of Proto-Germanic*. Leiden; Boston: Brill.
- LEW = Fraenkel, E. (1955-65). *Litauisches etymologisches Wörterbuch*. Heidelberg; Göttingen: Universitätsverlag Winter; Vandenhoeck & Ruprecht.
- Liden, E. (1899). "Ein baltisch-slavisches Anlautgesetz". *Göteborgs Högskolas Årsskrift*, 5.
- Machek, V. (1957). *Etymologický slovník jazyka českého a slovenského*. Praha: Nakladatelství Československé akademie věd.

- ME = Mühlenbach, K. (1923-25). *Lettisch-Deutsches Wörterbuch, ergänzt und fortgesetzt von J. Endzelin*. I-IV. Riga: Herausgegeben vom lettischen Bildungsministerium.
- LIV² = Rix, H.; Kümmel, M.; Zehnder, T.; Lipp, R.; Schirmer, B. (2001). *Lexikon des Indogermanischen Verben*. 2-te Aufl. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Orel, V. (1998). *Albanian Etymological Dictionary*. Leiden; Boston; Köln: Brill
- Ozoliņš, K. (2015). *Revisiting Proto-indo-european Schwebeablaut. A Dissertation Submitted in Partial Satisfaction of the Requirements for the Degree Doctor of Philosophy in Indo-European Studies*. Los Angeles: University of California.
- Persson, P. (1912). *Beiträge Zur Indogermanischen Wortforschung*. Bde 1-2. Uppsala: A.-B. Akademiska Bokhandeln.
- Rejzek, J. (2001). *Český etymologický slovník*. Voznice: Leda.
- Skok P. (1973). *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. Knj. I-IV. Zagreb: Jugoslavenska Akademija Znanosti i Umjetnosti.
- Smoczyński, W. (2021). *Słownik etymologiczny języka litewskiego. Wydanie drugie, poprawione i znacznie rozszerzone. Na prawach rękopisu*. <https://rromanes.org/pub/alii/Smoczy%C5%84ski%20W.%20S%C5%82ownik%20etymologiczny%20j%C4%99zyka%20litewskiego.pdf>.
- Steer, Th. (2015). *Amphikinese und Amphigenese: Morphologische und phonologische Untersuchungen zur Genese amphikinetischer Sekundärbildungen und zur internen Derivation im Indogermanischen*. Wiesbaden: Reichert Verlag. <https://doi.org/10.29091/9783954906475>.

Towards the Definition of the Bulgarian Word Order System

Anton Zimmerling

Pushkin State Russian Language Institute; Institute of Linguistics Russian Academy of Science, Russia

Abstract The world's languages with clitic clusters pattern with four major types of the word order systems, depending on the presence or absence of the 2P condition and the parameter of clitic-verb adjacency. Bulgarian has a double-focus system of clitic placement (2P condition and clitic-verb adjacency), which has typological parallels outside Europe but lacks direct counterparts in modern European languages. Neither the analogy with standard 2P languages without clitic-verb adjacency nor the analogy with the Romance systems with vP-internal clitics captures the profile of the Bulgarian clitic syntax. Historically, the rise of the clitic-verb adjacency is an innovation of Bulgarian, but its exact time and triggers are unclear. The language of the Wallachian letters (ca. 1386-1509) has a marked tendency towards the clitic-verb adjacency and is typologically similar to Modern Bulgarian but still has residual #XP – CL – [Y] – V orders. This idiom spoken by the L2 speakers of Middle Bulgarian cannot be viewed as a source of the Modern Bulgarian but hints that the clitic-verb adjacency parameter could develop in the history of Bulgarian because of the contact influence on the part of some Non-Slavic clitic systems.

Keywords Typology. Word order. Clitics. Clitic clusters. Bulgarian. Syntax. Prosody.

Summary 1 Ordering of Non-clitic Categories. – 1.1 Greenbergian Typology. – 1.2 Sentence Cartography. – 2 Clitics and Word Order Typology. – 2.1 Wackernagel's Law and Tobler-Mussafia's Law. – 2.2 Clitic Hosts and Clausal Architecture. – 2.3 Clitic Clusters and Clustering Clitics. – 3 Word Order Systems with Clustering Clitics. – 4 Defining the Bulgarian Word Order System. – 4.1 Synchrony. – 4.2 Diachrony and the Inventory of Clitics. – 4.3 Old Church Slavonic and Vernacular Texts. – 5 Middle Bulgarian: The Wallachian Letters. – 6 Conclusions.



Peer review

Submitted	2021-10-13
Accepted	2021-12-22
Published	2022-02-15

Open access

© 2021 | Creative Commons Attribution 4.0 International Public License



Citation Zimmerling, A. (2021). "Towards the Definition of the Bulgarian Word Order System". *Balcania et Slavia*, 1(2), 239-256.

1 Ordering of Non-clitic Categories

Word order systems can be classified in terms of linearisation principles. For the ordering of non-clitic categories, two different measures – branching conditions and sentence cartography – are available.

1.1 Greenbergian Typology

The approach introduced by Greenberg (1963) is based on branching conditions, i.e. relative ordering of the head (X^0) vs complement in binary groups (A^0 , B) and the so-called sentence formula or basic word order showing the ordering of three diagnostic sentence categories – subject (S), object (O) and verb/predicate (V): in the pair (V^0 , O), the verb is analysed as the head; in the pair (S^0 , V), the subject under standard assumptions is analysed as the head. In the early versions of the Greenbergian typology (cf. Hawkins 1983), all languages including the so-called scrambling (free word order) languages, where the phrasal categories can be ordered in more than one way, have the basic word order per definition. With this approach, Slavic languages are relatively uniform: they can be classified as SVO languages with a predominant right branching. The modern data-oriented version of the Greenbergian analysis, based on the World Atlas of Language Structures (WALS) sample of 1376 world's languages (Dryer 2013b), confirms the existence of all six logically possible combinations of S , V , and O , i.e. SOV, SVO, VSO, VOS, OVS, OSV, although the last three types are rare (2,9% of the sample). The seventh class in Dryer's classification – languages lacking a dominant word order (189 languages, 13,73%) – is rather a concession to the agnostic view than a challenge to the initial Greenberg's assumption that all languages have a basic word order. I interpret it as a technical requirement that all languages in a typological sample must be described at the same depth of analysis: if a basic word order in a European language like German, Dutch, Frisian, Hungarian, Modern Greek, Welsh Romani, or Belorussian can only be established based on a formal syntactic analysis, which has not been done for other languages from the sample, the language patterns with the class lacking a basic word order.¹ A further problem with the input data is that typological databases occasionally put languages with similar word order systems into different classes because of the decisions made by the authors of the ref-

I am grateful to the anonymous reviewers for the valuable comments. The responsibility for all shortcomings is on the Author.

1 The notion of 'depth' is discussed in Haspelmath 2019.

erence grammars. For example, Modern Belorussian is classified in Mayo (1993, 924) as a language lacking a basic word order and allowing both VO and OV,² while Modern Russian is classified as an SVO, VO language (Bivon 1971) in a description made 22 years earlier. It is unclear whether such decisions are motivated by the real differences between the languages, by the theoretical sympathies of the authors, or by the changing beliefs in the linguistic mainstream.³

1.2 Sentence Cartography

Another approach for ordering sentence categories has been known as ‘template analysis’ or ‘sentence cartography’. The insight behind it is that there is a correlation between word classes/types of category and their linear position so that each category has its diagnostic slot in the template (‘map’) of a sentence. Sentence templates can be interpreted both as language-specific rules and as a manifestation of universal ordering principles. Early versions of the template analysis arose as empiric generalisations on the syntax of selected Germanic languages with a presumably fixed order of the post-finite categories (Diderichsen 1946), while classical versions of the 1990s-2000s combine the template analysis with the hypothesis on generalised phrase structure (Cinque 1999; 2014; Belletti 2004). The cartographic approach is compatible with the analysis of the word order variation. However, if a category alternates between two or more slots P_1, P_2, \dots, P_n , sentence templates become bulky since the alternations add extra slots, cf. the discussion in Zimmerling (2013, 28-33). Therefore, cartography is generally associated with those zones in clausal structure, where the ordering is more rigid, e.g. the clausal left periphery (Krapova 2002; Benincà, Munaro 2011), or clitic-internal ordering.

2 Clitics and Word Order Typology

The term ‘clitic’ emerged as a generalisation of the ancient terms *enclitic* ($X=CL$) and *proclitic* ($CL=X$), i.e. weak stress elements adjoining to their prosodic hosts (stressed words) from the right or the left, respectively (Spencer, Lúis 2010). It has been introduced to language

² The WALs sample for the V^0 , O languages includes 1516 languages, 101 of which (6,59%) are classified as “languages with both VO and OV” (Dryer 2013c).

³ As for Modern Bulgarian, all sources cited in Dryer 2013b and 2013c. Cf. Hubenova et. al. 1968, 31; Scatton 1984, 374; Dyer 1992, 16, 55, 149, who classify it as SVO, VO language.

theory by Arnold Zwicky (1977), who claimed that prosodically deficient elements incapable of making a phonetic word without combining with other elements (= ‘phonetic clitics’) are also syntactically deficient, i.e. have a special distribution in syntax and take positions not available for the non-clitic categories.⁴

2.1 Wackernagel’s Law and Tobler-Mussafia’s Law

The first observations that clitics constrain the clausal syntax were made in the late nineteenth century by Berthold Delbrück and Jacob Wackernagel. The latter showed that in some Old Indo-European languages including Old Greek, Old Indian, Avestan, Old Persian fixed-position enclitics took the clausal-second position after the first stressed word and proved that this feature of the second-position clitics (2P clitics) is triggered by a general principle, which does not depend on the syntactic type or the origin of the clitic element (Wackernagel 1892). The discovery of Hittite and Luwian confirmed that Wackernagel’s law was characteristic of most Old Indo-European languages. A definitional property of 2P clitics is that they attach to at least two different syntactic types of the host (X, Y), but do not license them simultaneously:

(i) #X=C, #Y=CL, *#X, Y-CL

The so-called Tobler-Mussafia’s law (Tobler 1875; Mussafia 1886) is based on the distribution of Old French and Old Italian clitic pronouns: they are verb-adjacent and license the orders V=CL ~ CL=V clause-internally, but not clause-initially, *#CL-V. This is a consequence of Wackernagel’s law, preventing strict enclitics from taking the clause-initial position. Similar conditions are attested in Old Spanish (Wanner 1996). Bulgarian predicate clitics (clustering pronouns and auxiliaries) have the same bundle of features: they are both verb-adjacent and clause-internal. For such languages, it is possible to treat Tobler-Mussafia’s law as a non-syntactic condition imposed on the parameter of the verb-clitic adjacency (cf. Franks 2008). This conclusion is prompted by the fact that Old Romance word order systems are historically transitional from systems with Wackernagel’s law and 2P clitics to systems of the Modern Romance type with verb-adjacent clitics. Tobler-Mussafia’s law is redundant for the

⁴ The initial version of Zwicky’s analysis had a stipulation that special syntactic distribution is characteristic only of a subclass of clitics labelled “special clitics”, but not for the residue labelled “simple clitics”. However, simple clitics are elusive since their recognition crucially depends on the chosen theory of sentence derivation.

description of those languages with 2P clitics, which lack the parameter of obligatory clitic-verb adjacency and license configurations like $\#X=CL \dots V$.

2.2 Clitic Hosts and Clausal Architecture

Roman Jakobson (1935) was the first linguist who admitted that Wackernagel's law is instrumental in Modern Slavic languages, notably in the Balkan Slavic group, which partly retains the old accent conditions. Andrej Zaliznjak (1993) was the first linguist who proved that vernacular Old Russian had Wackernagel's law and explained, how the old East Slavic clitics were eliminated in the subsequent history of Russian (2008). Both Jakobson and Zaliznjak associated Wackernagel's law with 2P clitics after the first phonetic word and claimed that Modern Slavic languages placing 2P clitics after the first constituent (Slovenian, Bulgarian, Czech, Slovak, etc.) lost Wackernagel's law. This claim finds little support from linguistic typology since Wackernagel's law is not a unique feature of Old European languages. From the 1920s on, 2P clitics were discovered in many genetically not related world's languages from different areas, including Luiseño and Mayo (Uto-Aztecan), Ossetic, and Pashto (Indo-Iranian), Warlpiri and Djaru (Pama-Nyungan), etc. A part of these languages, e.g. Ossetic, Pashto, Kavineña (Tacanan), Kashibo-Kakataibo (Pano) places clausal 2P clitics after the first spelled-out constituent, not the first phonetic word, while some languages, e.g. Bosnian-Croatian-Serbian (BCS), Luiseño and Warlpiri license the first phonetic word/first constituent variation ($\#X/XP=CL$) (cf. Halpern 1996). Finally, clausal 2P clitics are also attested in languages that lack initial NPs and initial multi-word constituents, e.g. in Lummi (Salish), or Kabyle (Afroasiatic).

In subordinate clauses, 2P clitics generally attach to the complementizer (Comp=CL), though languages where the clitics skip the complementizer and attach to the next category Y (Comp Y=CL) are attested as well. The latter scenario is required in Pashto (2P, no verb-clitic adjacency), Tagalog (2P & verb-clitic adjacency) and optional in colloquial Czech (2P, no verb-clitic adjacency) and Bulgarian (2P & verb-clitic adjacency) (cf. Zimmerling 2013, 64, 114, 437). Additional types of the clitic hosts licensed by some languages with 2P clitics include initial proclitics and sentential complements ($[_{CP} \dots]=CL$): the latter option is regularly used in Cavineña, Kashibo-Kakataibo, and Slovenian, sporadically also in Czech, Upper Sorbian, South-Eastern Tepehuan, BCS, and Bulgarian (Zimmerling 2013, 154-62, 434-8). The principle of 2P clitic placement is strikingly uniform across the world's languages and compatible with different morphosyntax, while the mechanisms licensing diverse clitic hosts in

languages with 2P depend on the clausal architecture in the corresponding language. This makes the 2P syntax an interface phenomenon: a linguist must both explain the locus of the clitics, and the choice and internal structure of the hosts. The terms ‘prosodic’ and ‘syntactic’ used in two competing descriptions of the same data – the situation in BCS, a language with 2P clitics and the $\#X=CL \sim XP=CL$ variation (Radanović-Kocić 1996; Progovac 1996) – are misleading: they rather refer to different models of the prosody-syntax interface than to the possibility to eliminate all syntactic vs prosodic information from the description.

2.3 Clitic Clusters and Clustering Clitics

A cluster is a complex object satisfying the following criterion:

- (ii) If a syntactic domain contains two (or more) clitics *a* and *b* and they stay contiguously, they are placed in a rigid order $a > b$ (*a* immediately precedes *b*’).

The clitics conforming to the criterion (i) are called *clustering*, while the rules predicting clitic-internal ordering are called *clitic templates* or *ranking rules* (Zaliznjak 1993, 282; Franks, King 2000). True clusters must be distinguished from occasional sequences of phonetically adjacent clitics. The latter can belong to different hierarchically independent syntactic domains, while clustering clitics represent one domain.⁵ On the synchronic level, clitic templates are idiosyncratic since the template order cannot be derived from the order of non-clitic categories. However, it is up to a certain extent possible to explain it historically (Zimmerling 2012) and to reconstruct the principles of the prosody-to-syntax interface underlying the templates (Billings, Konopasky 2002). For Slavic languages, it has been shown that despite the clusters often contain clitics of more than one type – pronouns vs auxiliaries vs particles – no Slavic language licenses the insertion of an auxiliary into a block of clitic pronouns or vice versa. Other factors such as cliticization time or prosodic weight do not override in the Slavic languages the principle of grouping the clustering elements into blocks of clitics sharing the same syntactic category (Franks 2008; Zimmerling, Kosta 2013).

Clustering clitics are attested both in 2P languages and in languages with vP/VP-internal clitics. At the same time, there exist lan-

⁵ In languages with the parameter of clitic climbing, e.g. in Slovak and BCS, the clitic template of the matrix clause includes slots for the climbed clitics generated in the embedded clause (Zimmerling 2013, 137).

guages with single non-clustering 2P clitics. The ability to form a cluster is not an inherent characteristic of clitics: a clitic morpheme can be clustering in some configurations and non-clustering in other configurations in the same language. In Slavic languages, only clause-level clitics cluster. For example, Old Church Slavonic, Old Russian, and BCS accusative clitics are clustering as clause-level argument clitics but non-clustering as PP-level clitics, when they attach to prepositions. Old Russian free clitic же ‘emphatic marker’ was clustering as an emphatic sentential particle, but non-clustering as a non-sentential additive marker ($X \text{ же } Y$ ‘X and Y’), or as an identifier (Zaliznjak 1993, 281) (for the external parallels from Non-European languages, cf. Zimmerling 2012). The same holds for Bulgarian dative clitics: they are clustering as clause-level argument clitics but non-clustering as possessive DP-level clitics.⁶ Such facts support the claim that clusterization takes place in dedicated syntactic positions and is blocked in other positions. A possible solution to this puzzle is to analyse the clusters as phrases of a special kind, CLP (Zimmerling, Kosta 2013). A single clustering clitic is interpreted as an occurrence of CLP, while the number of clitics in a sentence depends on such factors as the valency frame of the predicate, the type of the clause, etc.

Some authors argued that all parts of the cluster must be prosodically homogeneous (Franks 2008, 95), but there is empirical counter-evidence against this claim. Clustering and non-clustering clitics have non-identical properties. Clusters arise as rhythmic patterns and their elements often represent different layers of cliticization, therefore some clustering elements can lack full properties of phonetic clitics. This was arguably the case with 1-2 p. present tense indicative BE-auxiliaries in Old Russian (Zaliznjak 2008, 37). Nevertheless, the ordering of these elements was rigid (225-7).

3 Word Order Systems with Clustering Clitics

Languages with clustering clause-level clitics pattern with four types depending on the presence/absence of the clitic-verb adjacency and the 2P principle constraining either the position of the clustering clitics (CL) or the position of the verb (V). Henceforth, I apply the notation introduced in Zimmerling 2013 and Zimmerling, Kosta 2013. The position of the pivotal category (CL or V) respective the clausal left border (#) is defined in terms of surface syntax, while the preceding 2P is analysed as X/XP, i.e. intra-clausal position licensing at least two

⁶ It is dubious that the suffixed article and the possessive pronoun in Bulgarian DPs like [_{DP} [книга-та]=му] represent the same level of cliticization.

different categories A v B, but not simultaneously (Zimmerling 2015).

iii $[_{XP} A \vee B] - 2P, *[_{XP} A \& B] - 2P$

The tag ‘W-system’ (‘W’ reads either ‘word’ or ‘Wackernagel’) stands for word systems with 2P clitics lacking the parameter of clitic-verb adjacency. W-systems are subdivided into W_1 -systems, i.e. languages lacking clause-initial NPs and W_2 -systems licensing them. W_2 -systems are further subdivided into W_{2A} -systems, where 2P clitics are placed after the first phonetic word and split the initial multi-word constituent ($\#[_{XP} W^1 = CL, W^2..W^n]$), W_{2B} -systems, where 2P clitics are placed after the first spelled-out constituent and W_{2C} -systems, which license the $X/XP = CL$ variation. W_2 -systems are attested in different areas and are common, while W_1 -systems are only attested in those verb-initial languages, where V1 alternates with an operator category (clause-initial quantifier or T(ense) A(spect) M(ood) marker).

The tag ‘W⁺-system’ (‘W⁺’ reads ‘modified W’)⁷ stands for word order systems with clitic-verb adjacency and the 2P condition defined for one of the pivotal categories – either CL or V. There are two main varieties. In the Philippine-type W⁺-systems, the pivotal category is CL, which takes clausal 2P. If clitics are present, the verb adjoins to them either from the left or from the right ($\#V-CL \sim \#XP-CL-V$). Word order systems of this subtype are characteristic of Austronesian languages, especially from the Central Philippine area. Bulgarian seems to be the only representative of this subtype in Modern Europe. In the Germanic-type W⁺-systems, the pivotal category is V, which takes the V2 or the V1/V2 positions, while the clitics are realised in the postverbal positions as CL2/CL3 depending on the position of the verb. Word order systems of this type are attested only in Old Germanic languages: the V2 and V1/V2 constraints are typologically rare outside Europe.

The tag W*-system (‘W*’ reads ‘degraded W’) stands for languages, where the 2P placement remains an option for the clause-level clitics and/or is not generalised for all types of clauses. This is, for example, attested in Old Church Slavonic, Modern Polish, Gurindi, and Mudburra (both are Pama-Nyungan languages).

Finally, the tag V-system (‘V’ reads ‘verb’) stands for languages of the Modern Romance type, where neither V nor CL has a fixed position respective to the clausal left border in the diagnostic type of clauses, but the verb and the clitics are adjacent ($\#...[V + CL]...$). V-systems constitute a majority of European word order systems with the clustering clitics, but exact parallels outside Europe are rare.

⁷ The term ‘W⁺-system’ was first coined by the author of this article in 2002.

Table 1 Word order systems with clustering clitics, after Zimmerling 2013

	W-systems				W ⁺ -systems		W ⁺ -systems	V-systems
	W ₁	W _{2A}	W ₂ W _{2B}	W _{2C}	Philippine type	Germanic type		
clitic-verb adjacency	no	no	no	no	yes	yes	no	yes
2P condition for CL	yes	yes	yes	yes	yes	no	(yes)	no
2P condition for V	no	no	no	no	no	yes	(no)	no
Languages	Lummi, Kabyle	Hittite, Old Greek, Old Novgorod Russian	Czech, Slovak, Slovenian, Ossetic, Pashto, Cavineña	BCS, Luiseño, Warlpiri	Tagalog, Cebuano, Bikol, Masbatenyo, Bulgarian, Old Italian	Old Icelandic, Middle Norwegian	Old Church Slavonic, Polish, Gurindi, Mudburra	French, Italian, Spanish, Rumanian, Albanian, Modern Greek

W-systems, W⁺-systems, and V-systems represent stable ordering principles, while W⁺-systems are transitional, which partly explains the uneven coverage of these word order systems.

4 Defining the Bulgarian Word Order System

4.1 Synchrony

There are two main approaches to defining the Bulgarian word order system – to analyse it as a language with 2P clitics with an added CL-V adjacency (Gălăbov 1950; Dimitrova-Vulchanova 1999) or as a language with verb-adjacent clitics with a non-syntactic 2P condition (Tobler-Mussafia's law) imposed on the distribution of CL (Rå Hauge 1976; Franks 2008). The first solution is more popular in functional studies, the second one in formal descriptions of Bulgarian. Both approaches are complementary, but from the typological perspective, the first one is preferable since the features '+ 2P' and '+ verb-adjacent' do not exclude each other. The parallel between Bulgarian and Central Philippine languages was first acknowledged in Billings, Konopasky 2002, cf. also a survey of Philippine W⁺-systems in Lee, Billings 2005. The dichotomy of mutually exclusive taxons '2P clitics' vs 'verb-adjacent clitics' is prompted by the current distribution of W-systems and V-systems in modern Europe, whereas Bulgarian is the sole representative of the Philippine-type W⁺-system in this area. V-systems of the Modern Romance type, where V and CL are

adjacent, but neither V nor CL has a fixed position respective to the clausal left border, are rare outside Europe.

The orientation of Bulgarian clitics towards the left periphery is proved by such tests as the impossibility of the *NP1-NP2-V-CL order, cf. the pair of sentences, where both nominal arguments are indexed by the pronominal clitics, but the *#X_i-Y_j V-[_{CL} a_i b_j] order is blocked:

- (1) Blg. a. [На Иван]_i [книгата]_j = сѣм = му_i = я_j върнал
'I have returned the book to Ivan', lit. 'To Ivan, the book, I gave it back to him'.
b. *[На Иван] [книгата] върнал = сѣм = му = я.

Another test confirming the 2P-orientation of Bulgarian clustering clitics is based on the distribution of periphrastic verbal forms containing an optative marker (not a clustering clitic itself). These forms host the clitics, but do not combine with other elements in XP:

- (2) Blg. a. [купил бих] (1) = я_i книгата_i
'I would buy that book'.
b. [книгата]_i (1) = я_i бих купил
'I would by that book'.
c. *[книгата]_i (1) [купил бих] (2) = я_i.

4.2 Diachrony and the Inventory of Clitics

The exact definition of the Proto-Slavic clitic system is a matter of debate. Jakobson (1935) and Gălăbov (1950) assumed that Proto-Slavic had Wackernagel's law, i.e. it was a standard W-system in terms of this paper, more exactly a W_{2A}-system. This assumption is based on the fact that the core of the 2P clitic inventory – the particles же 'emphatic marker', ли 'yes-no marker' and the 1-2 p. dative pronouns ми '1Sg.Dat', ти '2Sg.Dat', си 'Refl.Dat', ны '1Pl.Dat|Acc', вы '2Pl.Dat|Acc', на '1Du.Dat|Acc', ва '2Du.Dat|Acc' – is common Slavic, and all Old Slavic languages use them as clustering 2P elements. The same holds for the 2P particle бо 'causal marker' preserved in South Slavic and Southern Old Russian (Zaliznjak 2008, 31). The short accusative forms ма '1Sg.Acc', та '1Sg.Acc', са 'Refl.Acc', и '3Sg.M.Acc', ю '3Sg.F.Acc', ю '3Sg.M.Acc', ъ '3Sg.N.Acc', я '3Du.M.Nom|Acc' have common Slavic morphology and cluster clause-internally. However, they retain some non-clitic side uses in Old Church Slavonic texts, which shows that they were cliticised later than the 1-2 p. dative pronouns (Vaillant 1948, § 262; Zaliznjak 2008, 36).

In more recent times, the idea that Bulgarian inherited the 2P condition from Proto-Slavic was challenged. Pancheva (2005) argues that

not only the rise of the CL-V adjacency but also the uniform treatment of clitic pronouns and BE-auxiliaries as 2P clitics was an innovation realised during the history of Bulgarian. It is well-known that Slavic present tense indicative BE-auxiliaries represent a relatively recent layer of cliticization and were made part of the cluster later than dative and accusative pronouns. This is mentioned by Zaliznjak (1993, 285) regarding the positions of Old Russian 1-2 p. auxiliaries, which show the enclitic behaviour only as part of the perfect construction with the *-l* participle. The slots of auxiliary clitics in a cluster are different in three areal types of the Slavic template. Old East Slavic dialects place the 1-2 p. auxiliary clitics⁸ after the block of clitic pronouns in the slot labelled 'AUX2' in Zimmerling 2013, 110, 337.⁹

(iv) East Slavic [_{CLP} [_{PTCL} ...] [_{PRON} ...] AUX2]

West Slavic languages place the auxiliary clitics before the block of clitic pronouns in the slot labeled AUX1 (Zimmerling 2013, 110, 337).

(v) West Slavic [_{CLP} [_{PTCL} ...] AUX1 [_{PRON} ...]]

Finally, Balkan Slavic languages including Bulgarian use both AUX1 and AUX2: most auxiliaries take AUX1, while the 3Sg =*je* takes AUX2 (Franks, King 2000; Zimmerling 2013, 111, 337).

(vi) Balkan Slavic [_{CLP} [_{PTCL} ...] AUX1 [_{PRON} ...] AUX2]

It is plausible that the AUX1 slot is an innovation compared to AUX2: vernacular Old East Slavic texts with clitic clusters and the active AUX2 slot date back to the eleventh century, while the earliest comparable West Slavic texts come from the first half of the fourteenth century. The migration of auxiliary clitics from AUX2 to AUX1 is attested in the fourteenth century in Old Serbian (Pavlović 2011) and Old Czech (Zimmerling 2013, 474-8). However, neither the observed geographical variation of the AUX slots nor the inconsistent placement of the auxiliaries in some Old Slavic dialects proves that Proto-Slavic lacked Wackernagel's law. Wackernagel's law predicts the 2P placement of all elements that belong to the class of clustering clitic-

⁸ Most Old Russian dialects lack overt 3 p. present tense auxiliary clitics. This holds for the North-Western (Novgorod and Pskov) and likely for the Western (Smolensk and Polotsk) dialects.

⁹ A historical continuation of the East Slavic template rule is represented in Carpathian Rusyn dialects (Tolstaya 2012). These dialects have extra slots for the new plus perfect auxiliaries.

ics. If this class is expanded during a period of time, it means that the same ordering principle is applied to a larger inventory of clitics. New clitics add extra slots to the existing templates or fill in the existing slots, but neither of these scenarios brings about a change of the word order system.

4.3 Old Church Slavonic and Vernacular Texts

Clitic clusters represent rhythmic patterns. One can disregard the type of the text only in a situation, when such patterns and word orders with clusters are generalised in all forms of a language, e.g. in the present-day codified Slavic languages. The situation in Old Slavic languages is less evident. Wackernagel's law and the conditions of standard W_2 -systems are best displayed in the texts close to the oral speech, like birch bark letters and the direct speech fragments in the Old Russian languages. This has been proved in detail by Zaliznjak (2008), who demonstrated that the inconsistent bookish Old Russian W^* -system was derived from the consistent colloquial Old Russian W -system by a parametric change licensing the end position of the clitics and the choice of the hosts. This change of parameters has both a diachronic and a stylistic dimension: the authors could shift the register in some hybrid text genres, e.g. in Old Russian chronicles. Letters and diplomas written by the professional scribes are generally a reliable source, too, especially in a situation, where non-bookish vs bookish texts require different forms of a language if not different languages (Church Slavonic vs vernacular Slavic idiom). Meanwhile, Old Church Slavonic as liturgic language poses a problem even for the group of Old South Slavic dialects most close to it: in those genres, where Old Church Slavonic was the required standard, the authors modelled their word order after the precedent texts, not their everyday speech. The same presumably holds for the early Old Church Slavonic, since the translators of the Slavic Bible in many cases copied the word order and clitic positions from the Greek text (Vaillant 1948, 261).

5 Middle Bulgarian: The Wallachian Letters

The word order system of the Wallachian letters (ca. 1386-1509) is similar to the Modern Bulgarian W^+ -system, though it cannot be viewed as its direct source: these letters are written by the L2 speakers of Middle Bulgarian, who also spoke a form of Rumanian and probably other languages. The clustering clitics take 2P, the contact orders of CL and V prevail. Still, there are 51 examples of the order $\#XP\text{-}CL\text{-}[Y]\text{-}V$, with non-clitic categories intervening between

CL and V.

- (3) Middle ѿ сем же віе добрѣ знаете, како =ви =смы до нинѣ¹ азъ² защитил
Blg. ѿ Турци, [LVI: Vlad I Dracul. (c. 1431-1446)].¹⁰
‘And you know it for sure, how I² have defended you from the Turks till
now¹’.

Examples like (3) allow to qualify the language of Wallachian letters as a W-system and not as a W⁺-system like Modern Bulgarian or Tagalog. Another deviation from Modern Bulgarian is that some authors licensed the clustering clitics clause-initially. This is attested a dozen times in letters nos. 74, 179, 212, 213, 215, 216, 226: in 3 examples the CL1 order comes up in yes-no questions like (4), in 8 examples it is attested after the initial proclitic и ‘and’, cf. (5) and in one example, cf. (6) the clitic is likely used after a prosodic break.

- (4) Middle Тои не знам: е=ли=ви доишла книга въ рѣце, или не е. [CCXII:
Blg. Petru Uroș, trimesul Brașovenilor la împăratul Sigismund, și Ștefan,
logofătul lui Vlad Dracul. (c. 1432-1437)].¹¹
‘I am not aware of that: did the book reach your hands, or not.’
- (5) Middle Казахъ=ми ѿни: ѿста Мартин, да доде съ друуги товар. И=ми=
Blg. поръчи съ Ханеш. ако ест хтѣніе ти, пусти ми съ Ханеш други
дукати, понеже ест zde и ѿщъ мѣд, да я купѣ. [CCXIII: Gherghe
Lascar, cămărașul lui Vlad I Dracul. (c. 1437)].¹²

¹⁰ An anonymous reviewer indicates that word orders like (3) are marginally possible with some adverbials in some styles of Modern Bulgarian.

¹¹ The Wallachian Bulgarian example (4) likely has contrastive stress on the initial auxiliary *e*. In this case, the syntax of (4) is compatible with Modern Bulgarian sentences like Blg. Сте= ли или не=сте начело на движение за съединение на Княжеството с Източна Румелия? I am grateful for the anonymous reviewer for providing this example and for the discussion. The fronting of the auxiliary clitic in contexts, where it gets contrastive stress can be explained in a twofold way: a) prosodically – the deficient elements that are used as enclitics clause-internally are boosted and ‘repaired’ in contrastive contexts, b) information-structurally – the clustering clitics that normally take 2P, are fronted if the 1P order has communicative motivation (Zimmerling 2015). It should be mentioned that standard definitions of clitics generally specify that true clitics cannot take contrastive stress, be negated or conjoined. While this condition holds for the majority of word order systems with clustering clitics, some languages have an option for the early placement of clitics in certain non-canonic positions.

¹² An anonymous reviewer points out that similar examples, where the clustering clitics occur in such initial or quasi-initial positions can be found in Modern Bulgarian texts as well, cf. Blg. И той се=нахвърля върху ми. А ми= е жал за момчето! Given that the variant ⁷А жал= ми=е за момчето! is rated as odd, these facts suggest that Modern Bulgarian and the language of Wallachian Bulgarian letters have an option for the emphatic fronting of (some) clustering clitics. Similar type of clitic fronting is attested in Slovenian, which is a W_{2B} – system without the clitic-verb adjacency. The default order

‘They told me that Martin stayed there and is going to come with more goods. **And** send me a message with Hanes: “If you like, send me more ducats with Hanes, because there is more honey here so that I buy it”’.

- (6) Middle Друго, **ми**=дръжѣтъ Кръстѣ Рошѣул ѡнеи книге, не щет да=**ми**=**их**
Blg. даст; [CXXVI: Dragomir Udriște, vornicul lui Vlad III Călugărul. (c. 1482-1492)].
‘Second, Krstia Roshul keeps these books **from me** and is not going to give *them to me*’.

The authors of the examples (4)-(6) likely lacked the Tobler-Musaffia’s law in their idioms, but there are no sufficient grounds to project this feature to Middle Bulgarian in general. The dialectal variation in the corpus of the Wallachian letters is not unlikely. Uroș (1432-37), Gherghie Lascar (1437), and Dragomir Udriște (1482-1492) or their scribes could have CL1 in their Middle Bulgarian dialects and/or in their other language (e.g. Old Rumanian). However, the contexts of examples (4) and (5) are not entirely telling for the comparison with Modern Bulgarian since they involve such pragmatic mechanisms as contrastive stress on the perfect auxiliary *e* in (4) and emphasis in (5): in these contexts, Modern Bulgarian occasionally or regularly licenses contrastive and emphatic fronting of the clustering clitics.¹³

Regarding the auxiliary slots, the language of the Wallachian letters shows an intermediate-type clitic template like Old Czech or Old Serbian, with two slots – AUX2 and AUX1 – available for most auxiliary clitics. The complementary distribution of the 3Sg =*e* (AUX2) vs all other present tense indicative BE-clitics (AUX1) diagnostic for the Modern Balkan type of the template has not yet established. Moreover, AUX2 is the main slot, while AUX1 is a reserved option. Past tense auxiliaries opened an extra slot AUX4 to the right from AUX2, cf. example (7). Similar templates are attested in some Modern Carpathian Rusyn dialects (Tolstaya 2012).

- (7) Middle И по сих давам у знаніе господству=**ви** израді ѡвою чловѣкъ мою
Blg. на име Стѣнислав, како=**му**=**ест**=**бил** дал¹ Марко един синъ=**му**, та да=**га** научит² ѡзик, да=**му** плащат; а ѡн, кѣда=**ест** бил³ летоска, ѡн=**ест** бежал⁴, тако=**ест** ѡтишел⁵ през Дунав и Хръсова та=**се**=**ест** потурчил⁶. [CCXLVII: Carstian, vornicul lui Radu IV cel Mare. (c. 1496-1507)].

for Slovenian clustering clitics is 2P, cf. Slv. *Videl=sem=ga* ‘I have seen it’. However, colloquial Slovenian has an additional 1P option, when the clustering clitics are fronted in verificational and emphatic contexts, cf. Slv. *#Sem=ga=videl* ‘I have indeed seen it’.

¹³ See the details in footnotes 11 and 12. I am grateful to the anonymous reviewer for the valuable comments.

‘And now I notify your grace about my man, by the name Stānislav. Marko had previously sent¹ him one of his sons, so that Stānislav would teach² him the language for payment. But he ran⁴ away from where he was³ last summer, so he drove down⁵ the Danube from Hāršovo/Hāršova and converted⁶ to Islam [lit. became⁶ a Turk]’.

(7’)

И	по сих	давам	у	знание	господству	=ви	
And	till now	gave:1SG	by	knowledge	dominion	you: 2PL	
израдї	ѡвою	чловѣкъ	мою	на	име	Стѣнислав,	
sake	that	man	mine	by	name	Stānislav	
како	=му	=ест	=бил	дал ¹	Марко		
how	him:CL.DAT	is: CL.3SG	was: 3SG	gave: 3SG	Marco		
един син	=му	та	да	=га	научит ²	язик,	
one son	him: CL	so	to	him: CL.ACC	teach	language	
да	=му	плащат;	а	ѡн,	къда	=ест	бил ³
to	him: DAT.CL	pay: 3PL	but	he	when	is: CL.3SG	was: 3SG
летоска,	ѡн,	=ест	бежал, ⁴	тако	=ест	ѡтишел ⁵	
last.summer	he	is: CL.3SG	ran: 3SG	so	is: CL.3SG	departed: 3SG	
ПРЕЗ	Дунав и	Хрѣсова	та	=се	=ест	потурчил ⁶	
Along	Danube and	Hāršovo	so	REFL.CL is: CL.3SG		became.a.Turk: 3SG	

[CCXLVII: Carstian, vornicul lui Radu IV cel Mare. (c. 1496-1507)].

‘And now I notify your grace about my man, by the name Stānislav. Marko had previously sent¹ him one of his sons, so that Stānislav would teach² him the language for payment. But he ran⁴ away from where he was³ last summer, so he drove down⁵ the Danube from Hāršovo/Hāršova and converted⁶ to Islam [lit. became⁶ a Turk]’.

The Middle Bulgarian template is shown schematically in (vi).

(vii) Middle Bulgarian [_{CLP} [_{PTCL} ...] AUX1 [_{PRON} DAT ACC] AUX2 AUX4]

6 Conclusions

Bulgarian has a double-focus system of clitic placement (2P condition and clitic-verb adjacency), which is not typologically rare but lacks exact counterparts in modern European languages. Therefore, neither the analogy with standard 2P languages without clitic-verb adjacency nor the analogy with the Romance systems with vP-internal clitics captures the profile of the Bulgarian clitic syntax. Bulgarian inherited the principle of 2P clitic placement and part of the clustering clitics inventory from Proto-Slavic, while the clitic-verb adjacency is an innovation. Its exact time and triggers are obscure, but a contact influence of a non-Slavic Balkan word order system is probable. The language of the Wallachian letters has residual #XP-CL-[Y]-V orders but is otherwise typologically similar to Modern Bulgarian. However-

er, this idiom spoken by the L2 speakers of Middle Bulgarian cannot be viewed as a direct ancestor of Modern Bulgarian.

Bibliography

- Belletti, A. (ed.) (2004). *Structures and Beyond. The Cartography of Syntactic Structures*, vol. 3. New York: Oxford University Press.
- Benincà, P.; Munaro, N. (2011). *Mapping the Left Periphery: The Cartography of Syntactic Structures*. New York: Oxford University Press.
- Billings, L.; Konopasky, A. (2002). "The role of morphology in ordering verb-adjacent clitics: from syntax to prosody in Bulgarian and Tagalog". Alexiadou, A. et al. (eds), *Linguistics in Potsdam*, vol. 19. Potsdam: Institut für Linguistik, Universität Potsdam, 1-26.
- Bivon, R. (1971). *Element Order*. Cambridge: Cambridge University Press. Studies in the Modern Russian Language 7.
- Bošković, Ž. (2004). "Clitic placement in South Slavic". *Journal of Slavic Linguistics*, 12(1-2), 39-90.
- Cinque, G. (1999). *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. New York: Oxford University Press.
- Cinque, G. (2014). "Again on Tense, Aspect, Mood Morpheme Order and the 'Mirror Principle', Functional Structure from Top to Toe". *The Cartography of Syntactic Structures*, vol. 9. Oxford: Oxford University Press, 232-65.
- Diderichsen, P. (1946). *Elementær Dansk Grammatik*. København: Gyldendal.
- Dimitrova-Vulchanova, M. (1999). "Clitics in the Slavic languages". van Riemsdijk, H. (ed.), *Clitics in the languages of Europe* (Eurotype 20-5). Berlin; New York: Mouton, 83-121.
- Dryer, M.S.; Haspelmath, M. (eds) (2013a). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Dryer, M.S. (2013b). "Order of Subject, Object and Verb". Dryer, Haspelmath 2013. <http://wals.info/chapter/81>.
- Dryer, M.S. (2013c). "Order of Object and Verb". Dryer, Haspelmath 2013. <http://wals.info/chapter/83>.
- Dyer, D. (1992). *Word Order in the Simple Bulgarian Sentence*. Amsterdam: Rodopi.
- Franks, S. (2008). "Clitic Placement, Prosody and the Bulgarian Verbal Complex". *Journal of Slavic linguistics*, 16(1), 91-137.
- Franks, S.; King, T.H. (2000). *A Handbook of Slavic Clitics*. New York: Oxford University Press.
- Gălăbov, I. (1950). "Zur Frage der bulgarischen Enklitika". *Zeitschrift für slavische Philologie*, 10(2), 417-18.
- Greenberg, J.H. (1963). "Some Universals of Grammar with Particular Reference to the Order of the Meaningful Elements". *Universals of Language*. Cambridge (MA): MIT Press, 73-113.
- Halpern, A.L. (1996). "Introduction". Halpern, Zwicky 1996, ix-xxiii.
- Halpern, A.L.; Zwicky, A.M. (eds) (1996). *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications.
- Haspelmath, M. (2019). "Ergativity and the Depth of Analysis". *Rhema. Рема*, 4, 118-30.
- Hawkins, J. (1983). *Word Order Universals*. New York: Academic Press.

- Hubenova, M.; Dzhumadanova, A.; Marinova, M. (1968). *A Course in Modern Bulgarian*, vol. 1. Columbus: Slavica.
- Jakobson, R. (1971). "Les enclitiques slaves". *Selected Writings*. Vol. 2, *Word and Language*. The Hague; Paris, 16-22.
- Krapova, I. (2002). "On the Left Periphery of the Bulgarian Sentence". *Working Papers*, vol. 12. Venezia: Università Ca' Foscari Venezia, 107-28.
- Lee, C.; Billings, L. (2005). "Wackernagel and Verb-Adjacent Clisis in Central Philippines". Heinz, J.; Ntelitheos, D. (eds), *Proceedings of the 12th meeting of the Austronesian Formal Linguistics Association (UCLA Working Papers in Linguistics 12)*. Los Angeles: Department of Linguistics, University of California, 241-54.
- Mayo, P. (1993). "Belorussian". Comrie, B.; Corbett, G.G. (eds), *The Slavonic Languages*. London: Routledge, 887-946.
- Mussafia, A. (1886). "Una particolarità sintattica della lingua italiana dei primi secoli". Avolio, C.; Biadene, L.; Ascoli, G.I. (a cura di), *In memoria Napoleone Caix e Ugo Angelo Canello: Miscellanea di filologia e linguistica*, 155-261. Firenze: Successori Le Monnier.
- Pancheva, R. (2005). "The Rise and Fall of Second-Position Clitics". *Natural Language and Linguistic Theory*, 23(1), 103-67.
- Pavlović, S. (2011). "Distribucija enklitika u starosrpskim poveljama i pismima XII i XIII veka" (The Placement of Enclitics in Old Serbian Documents of the XIIth and XIII Century). *Zbornik Matice srpske za filologiju i lingvistiku* (Matica Srpska Journal of Philology and Linguistics), LIV(2), 31-52.
- Progovac, L. (1996). "Clitics in Serbian/Croatian: Comp as the Second Position". Halpern, Zwicky 1996, 411-28.
- Radanović-Kocić, V. (1996). "The Placement of Serbo-Croatian Clitics: A Prosodic Approach". Halpern, Zwicky 1996, 429-48.
- Rå Hauge, K. (1976). "The Word Order of Predicate Clitics in Bulgarian". *Journal of Slavic Linguistics*, 7(1), 89-137. (Originally appeared as Meddelelser 10, Oslo: Slavisk-baltisk institutt, Universitetet i Oslo).
- Sadock, J.M. (1995). "A Multi-Hierarchy View of Clitics". *CLS 31*. Part 2, 258-79.
- Scatton, E.A. (1984). *A Reference Grammar of Modern Bulgarian*. Columbus: Slavica Publishers.
- Spencer, A.; Luís, A. (2012) *Clitics: An Introduction*. Cambridge: Cambridge University Press.
- Tobler, A. (1875). Review of Joules Le Coultre, *de l'ordre des mot dan Crestien de Troyes*. Extrait du programme de Pâques 1875 du Collège Vitzthum. Dresde 1875 ... [sic]. *Göttingische gelehrte Anzeigen*, 2(34), 1057-82. (Reprinted in his *Vermischte Beiträge zur französischen Grammatik*, vol. 5. Leipzig: Hirzel, 395-415).
- Tolstaya, M.N. Толстая М.Н. (2012). "Karpatoukrainskie enklitiki v južnoslavyanskoj perspective" Карпатоукраїнські енклітики в южнославянської перспективі (Carpathian Unkrainan Enclitics in a South Slavic Perspective). Карпато-балканський діалектний ландшафт; Язык и культура: 2009-2011 *Karpato-balkaniskij dialektnyj landshaft; Yazyk i kul'tura: 2009-2011* (Carpathian-Balkanian Dialectal Landscape: Language and Culture: 2009-2011), vol. 2. Moscow: Institute of Slavic Studies, 190-211.
- Vaillant, A. (1948). *Manuel du Vieux Slave*. Paris: Institut d'Études Slaves.
- Wackernagel, J. (1892). "Über ein Gesetz der indogermanischen Wortstellung". *Indogermanische Forschungen*, 1, 333-436.

- Wanner, D. (1996). "Second Position in Medieval Romance". Halpern, A.L.; Zwicky, A.M. (eds), *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications, 537-78.
- Zaliznjak, A.A. Зализняк, А.А. (1993). "Lingvistečeskie issledovanija" Лингвистические исследования (Linguistic Studies). Janin, V.L. Янин, В.Л.; Zaliznjak, A.A. Зализняк, А.А. (eds), *Novgorodskie gramoty na bereste. Iz raskopok 1984-1989* Новгородские грамоты на бересте (из раскопок 1984-1989 гг.) (Novgorod Birch Bark Letters from the Excavations 1984-1989). Moscow: Nauka, 191-344.
- Zaliznjak, A.A. Зализняк, А.А. (2008). *Drevnerusskie enklitiki* Древнерусские энклитики (Old Russian Enclitics). Moscow: Yazyki Slavyanskoj Kultury.
- Zimmerling, A. (2012). "A Unified Analysis of Clitic Clusters in the World's Languages". *Computational Linguistics and Intellectual Technologies*, 11(18), 726-38.
- Zimmerling, A. Циммерлинг, А. (2013). *Sistemy porjadka slov slavyanskix yazykov v tipologičeskom aspekte* Системы порядка слов славянских языков в типологическом аспекте (Slavic Word Order Systems in Typological Perspective). Moscow: Yazyki slavyanskoj kultury.
- Zimmerling, A. (2015). "Parametrizing Verb-Second Languages and Clitic-Second Languages". *WORLDCOMP' 15. Proceedings of the 2015th International Conference on Artificial Intelligence*, vol. 1. CSREA Press, 281-7.
- Zimmerling, A.; Kosta, P. (2013). "Slavic Clitics: A Typology". *STUF*, 66(2), 178-214.
- Zwicky, A.M. (1977). *On Clitics*. Bloomington: Indiana University Linguistics Club.

Rivista semestrale | Semestral journal

Dipartimento di Studi Linguistici
e Culturali Comparati
Università Ca' Foscari Venezia



Università
Ca'Foscari
Venezia